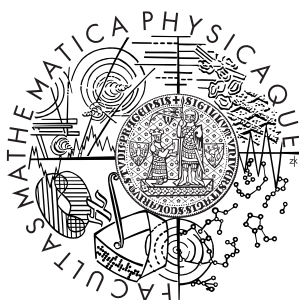


Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE



Dominika Šimjáková

Analýza finančních časových řad

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Jitka Zichová, Dr.

Studijní program: Matematika

Studijní obor: Finanční matematika

2007

Na tomto mieste by som sa rada poďakovala mojej vedúcej bakalárskej práce RNDr. Jitke Zichovej, Dr., za materiály a cenné rady, ktoré mi poskytla. Taktiež ďakujem za umožnenie práce s potrebným softwarom.

Prehlasujem, že som svoju bakalársku prácu napísala samostatne a výhradne s použitím citovaných prameňov. Súhlasím so zapožičiavaním práce a jej zverejňovaním.

V Praze dne 25. júna 2007

Dominika Šimjáková

Obsah

Úvod	5
Kapitola 1. Jednorozmerné rady	6
1.1. Exponenciálne vyrovňovanie	6
1.2. Boxova-Jenkinsova metodológia	7
1.3. Tri fázy tvorby modelu	15
1.4. Analýza dát	15
Kapitola 2. Mnohorozmerné rady	23
2.1. Základné vlastnosti mnohorozmerného náhodného procesu	23
2.2. Mnohorozmerné ARMA modely	23
2.3. Identifikácia ARMA modelu	25
2.3.1. Kovariančná a korelačná maticová funkcia	26
2.3.2. Parciálna autoregresná maticová funkcia	27
2.3.3. Parciálna korelačná maticová funkcia	28
2.4. Verifikácia modelu	29
2.5. Analýza dát	30
Záver	37
Literatúra	38

Název práce: Analýza finančních časových řad

Autor: Dominika Šimjáková

Katedra (ústav): Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Jitka Zichová, Dr.

e-mail vedoucího: zichova@karlin.mff.cuni.cz

Abstrakt: Práce sa zaoberá analýzou jednorozmerných a mnohorozmerných časových rád. Podrobne tu popisujem modely exponenciálneho vyrovnávania, Boxovu-Jenkinsovu metodológiu, AR a MA modely ako aj zmiešané ARMA modely. V praktickej časti bakalárskej práce spracovávam rady menových kurzov. Cieľom práce je nájsť vhodný model, ktorým bude možné spoľahlivo predpovedať budúci vývoj časovej rady. V jednorozmernom prípade sú rady spracovávané programom NCSS a v mnohorozmernom prípade softwarom Mathematica. Spracovávané dáta a výpis programu sú priložené na CD.

Klíčová slova: Analýza časových rád, mnohorozmerná časová rada, ARMA model

Title: The analysis of financial time series

Author: Dominika Šimjáková

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Jitka Zichová, Dr.

Supervisor's e-mail address: zichova@karlin.mff.cuni.cz

Abstract: The subject of the thesis is the analysis of univariate and multivariate time series. The models of exponential and Holt smoothing, AR and MA models as well as the ARMA models are described in detail. In the practical part of the bachelor work, I elaborate some time series of exchange rates. The aim of this work is to find an appropriate model, which would reliably forecast the future development of the series. In the one-dimensional case, the series are elaborated by the programme NCSS and in the multivariate case by the software Mathematica. The data and programme source code are enclosed on a CD.

Keywords: Time series analysis, multivariate time series, ARMA model

Úvod

Veľké množstvo dát vo financiách má podobu časovej rady, tj. hodnota rovnakého ukazovateľa sa zaznamenáva v rôznych konkrétnych časoch v minulosti.

V prípade mnohorozmerných rád sú uvažované vektory hodnôt. Úlohou je zistiť prípadnú závislosť medzi zložkami vektoru, zistiť závislosť na minulých pozorovaniach, identifikovať príslušný model, ktorý by tieto súvislosti popisoval a potom na jeho základe predpovedať dáta do budúcnosti.

Táto práca sa zaoberá aplikáciami jednorozmerných modelov a mnohorozmerných ARMA modelov na konkrétne finančné dáta. ARMA modely sú lineárne modely založené na predpoklade, že pozorovanie časovej rady v určitom čase je lineárnou funkciou predchádzajúcich pozorovaní a náhodnej zložky s nemenným rozdelením pravdepodobnosti.

Práca je členená do dvoch kapitol. Prvá kapitola je venovaná jednorozmerným radám. Sú tu popísané modely exponenciálneho vyrovnávania a ARMA modely. Kapitola je ukončená praktickou časťou, venovanou analýze menového kurzu. Druhá kapitola sa zaoberá teóriou mnohorozmerných ARMA modelov. Obsahuje definície kľúčových vlastností, popisuje fázy identifikácie modelu. Kapitola je ukončená praktickou časťou, venovanou rozboru dvojrozmernej rady menových kurzov.

Teoretická časť bola v prípade jednorozmerných rád spracovávaná podľa publikácie [1], v prípade mnohorozmerných rád podľa materiálov [2], [3].

Na priloženom CD sú súbory obsahujúce spracovávané dáta (dvojrozmkurzy2.txt) a použité programy (rada.nb, franky.S0).

KAPITOLA 1

Jednorozmerné rady

1.1. Exponenciálne vyrovňovanie

Exponenciálne vyrovňovanie patrí medzi adaptívne prístupy k trendovej zložke v časovej rade. V praxi je často používané. Medzi jeho prednosti patria malé pamäťové nároky a jednoduchá konštrukcia predpovedí. Jeho základy položili Holt a Brown-Meyer. Na rozdiel od metódy kĺzavých priemerov, pri ktorej sa snažíme vyrovňávať krátke úseky rady, o predom stanovenej dĺžke, polynomic-kymi krivkami, je výpočet vyrovnanej hodnoty pri exponenciálnom vyrovňovaní založený na všetkých dostupných pozorovaniach rady. Využíva sa metóda najmenších štvorcov, pričom sa váhy vyrovňovacej konštanty α v minimalizovanom súčte exponenciálne zmenšujú.

Minimalizujeme výraz tvaru:

$$(X_t - F_t)^2 + (X_{t-1} - F_{t-1})^2\alpha + (X_{t-2} - F_{t-2})^2\alpha^2 + \dots \quad 0 < \alpha < 1$$

OZNAČENIE. X_t predstavuje hodnoty pôvodnej rady, F_t predstavuje vyrovnané hodnoty.

POZNÁMKA. Predpokladáme, že vyrovňovaná rada má tvar $X_t = Tr_t + e_t$, kde Tr_t je trend časovej rady v čase t a e_t je chybová (náhodná) zložka v čase t .

Modelovanie trendovej zložky prebieha s použitím špeciálnych váhových funkcií, ktoré kladú väčší dôraz na časové periódy menej vzdialené do minulosti. Rovnica trendu sa v jednotlivých obdobiach mení, konštruje sa teda lokálny, nie globálny trend. V ďalšom výklade zmienime dva algoritmy používané pre rady s po častiach lineárnym trendom, ktoré sú implementované v programe NCSS.

OZNAČENIE. V podkapitole exponenciálne vyrovňovanie je zavedené rovnaké označenie ako v štatistickom programe NCSS.

Dvojité exponenciálne vyrovňovanie

Algoritmus pre konštrukciu vyrovnanej rady:

$$\begin{aligned} F_t &= a_t + b_t, \\ a_t &= X_t + (1 - \alpha)^2 e_{t-1}, \\ b_t &= b_{t-1} + \alpha^2 e_{t-1}, \\ e_t &= F_t - X_t. \end{aligned}$$

Predpoveď vykonaná v čase T pre hodnotu v čase $T + k$ má hodnotu $F_{T+k} = a_T + b_T k$. Voľbu parametru α a počiatočných hodnôt iteračného procesu prevádza štatistický software.

POZNÁMKA. Táto metóda je zahrnutá hlavne kvôli svojmu historickému významu. V praxi je preferovaný Holtov algoritmus.

Holtov lineárny trend

Algoritmus pre konštrukciu vyrovnanej rady a predpovedí je rovnaký ako v predchádzajúcom prípade, líši sa len spôsob výpočtu hodnôt a_t, b_t :

$$F_t = a_t + b_t,$$

$$a_t = \alpha X_t + (1 - \alpha)(a_{t-1} + b_{t-1}),$$

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1},$$

pre $0 < \alpha, \beta < 1$. Predpoveď sa konštruuje analogicky ako v predchádzajúcom prípade.

POZNÁMKA. Vyrovnávacie konštanty α a β môžu byť zadané objektívne alebo subjektívne. Subjektívne zadanie je založené na vlastnej skúsenosti a očakávaniach vývoja situácie. Pri objektívnom zadaní hľadáme optimálne hodnoty vyrovňovacích konštánt pre danú radu. To znamená, že sa musíme rozhodnúť, či budeme chcieť minimalizovať MSE, MAE alebo MAPE.

POZNÁMKA. Ocenenie kvality zkonštruovaných hodnôt vyrovnanej rady sa posudzuje na základe:

$$\text{MSE (Mean Squared Error)} \quad \frac{1}{n} \sum_{t=1}^n e_t^2$$

$$\text{MAE (Mean Absolute Error)} \quad \frac{1}{n} \sum_{t=1}^n |e_t|$$

$$\text{MAPE (Mean Absolute Percent Error)} \quad \frac{100}{n} \sum_{t=1}^n \left| \frac{e_t}{X_t} \right|$$

Štatistický software ponúkne užívateľovi optimálne hodnoty α, β založené na zvolenom minimalizačnom kritériu i počiatočné hodnoty a_0, b_0 .

1.2. Boxova-Jenkinsova metodológia

Boxove-Jenkinsove modely sú stochastické modely umožňujúce značnú flexibilitu v posudzovaní konkrétnej časovej rady. Umožňujú spracovávať rady, na ktoré klasická analýza časových rád nestačí. Napriek ich rozšírenému využitiu majú zopár nevýhod. Je nutné mať radu s minimálne 50 pozorovaniami. Praktická aplikácia je časovo aj finančne náročnejšia (je potrebný špeciálny štatistický software), strácame možnosť jednoduchšej interpretácie modelu.

Autokorelačné vlastnosti časových rád

Stacionarita (slabá) je vlastnosť, kedy príslušný stochastický proces má konštantnú strednú hodnotu, konštantný rozptyl a kovariančnú štruktúru invariantnú voči posunom v čase, tj.

$$\text{cov}(y_t, y_s) = \text{cov}(y_{t+h}, y_{s+h}) \quad \text{pre ľubovoľné } h.$$

Kovarianciu často normujeme na tzv. korelačný koeficient:

$$\text{cor}(u, v) = \frac{\text{cov}(u, v)}{\sqrt{\text{var}(u)\text{var}(v)}},$$

ktorý nadobúda hodnoty od -1 do 1 .

POZNÁMKA. V Boxovej Jenkinsovej metodológii je možné modelovať len stacionárne časové rady. Stacionaritu zaistíme pomocou vhodných transformácií (najčastejšie pomocou diferencovania).

Najpoužívanéjšie transformácie dát

- (1) **odpočítanie priemeru** $y_t - \bar{y}$ zaručí $Ey_t = 0$, ale nezaručí stacionaritu modelu, tzn., že transformácia je výhodná pri stacionárnych modeloch s $Ey_t \neq 0$
- (2) **odpočítanie trendu** $y_t - a - bt$ môže odstrániť nestacionaritu, transformácia dobre funguje pri lineárnom trende v časovej rade
- (3) **1. diferencia** $D_t = y_t - y_{t-1}$ výhodné pri nestacionarite v strednej hodnote
- (4) **2. diferencia** $D_t - D_{t-1}$ používa sa pri modeloch s nestacionaritou v rozptyle
- (5) **logaritmovanie** $\ln y_t$
- (6) **diferencie logaritmov** $\ln y_t - \ln y_{t-1}$ často používané na modelovanie finančných časových rád
- (7) **normované diferencie** $\frac{y_t - y_{t-1}}{y_t}$

POZNÁMKA. Transformácie dát uvedené pod číslami (3) a (4) sú niekedy označované ako modely ARIMA(p,d,q), kde číslo d označuje rád diferencie. Transformácie (6) a (7) znižujú rozsah kolísania okolo nulovej úrovne.

Autokovariančná a autokorelačná funkcia

DEFINÍCIA. Pre stacionárnu časovú radu y_t sa hodnota jej *autokovariančnej funkcie* γ_k v bode k definuje ako

$$\gamma_k = \text{cov}(y_t, y_{t+k}) = E(y_t - \mu)(y_{t+k} - \mu), \quad k = \dots, -1, 0, 1, \dots$$

a hodnota jej *autokorelačnej funkcie* ϱ_k v bode k ako

$$\varrho_k = \frac{\gamma_k}{\gamma_0} = \frac{\gamma_k}{\sigma_y^2}, \quad k = \dots, -1, 0, 1, \dots$$

kde $\mu = E(y_t)$ je stredná hodnota a $\sigma_y^2 = \gamma_0 = \text{var}(y_t)$ je rozptyl danej stacionárnej rady.

POZNÁMKA. Platí párnosť autokorelačných a autokovariančných funkcií ($\varrho_k = \varrho_{-k}$ a $\gamma_k = \gamma_{-k}$), takže pre ich popis sa stačí obmedziť na $k \geq 0$. Platí $\varrho_0 = 1$ a $|\varrho_k| \leq 1$. Grafický záznam ϱ_k sa nazýva *korelogram*.

Odhad autokovariančnej a autokorelačnej funkcie

Pre pozorovanú časovú radu y_1, y_2, \dots, y_n používame ako odhad strednej hodnoty:

$$\bar{y} = \sum_{t=1}^n \frac{y_t}{n},$$

odhad autokovariančnej funkcie:

$$c_k = \sum_{t=1}^{n-k} \frac{(y_t - \bar{y}) \cdot (y_{t+k} - \bar{y})}{n}, \quad k = 0, 1, \dots, n-1$$

odhad autokorelačnej funkcie:

$$r_k = \frac{c_k}{c_0}, \quad k = 0, 1, \dots, n-1$$

POZNÁMKA. Pre výpočet odhadov sa doporučuje $n > 50$ a $k < \frac{n}{4}$.

POZNÁMKA. Správanie autokorelačnej funkcie je dôležitým ukazovateľom, napovedá aký typ modelu je vhodné pre danú radu použiť. Hovoríme, že toto správanie *identifikuje* príslušný model. Pritom je dôležité určiť hodnotu $k = k_0$ (ak existuje), za ktorou začína byť autokorelačná funkcia nulová.

Parciálna autokorelačná funkcia

DEFINÍCIA. *Parciálna autokorelačná funkcia* ϱ_{kk} je definovaná ako parciálny korelačný koeficient y_t a y_{t+k} pri pevných hodnotách $y_{t+1}, \dots, y_{t+k-1}$.

Platí:

$$\varrho_{kk} = \frac{|\mathbf{P}_k^*|}{|\mathbf{P}_k|}$$

kde $||$ označuje determinant matice, \mathbf{P}_k je matica autokorelácií tvaru

$$\mathbf{P}_k = \begin{bmatrix} 1 & \varrho_1 & \dots & \varrho_{k-1} \\ \varrho_1 & 1 & \dots & \varrho_{k-2} \\ \dots & \dots & \dots & \vdots \\ \varrho_{k-1} & \varrho_{k-2} & \dots & 1 \end{bmatrix}$$

a matica \mathbf{P}_k^* , ktorá vznikne predefinovaním posledného stĺpca matice \mathbf{P}_k má tvar

$$\mathbf{P}_k^* = \begin{bmatrix} 1 & \varrho_1 & \dots & \varrho_1 \\ \varrho_1 & 1 & \dots & \varrho_2 \\ \dots & \dots & \dots & \vdots \\ \varrho_{k-1} & \varrho_{k-2} & \dots & \varrho_k \end{bmatrix}$$

POZNÁMKA. Platí $\varrho_{11} = \varrho_1$. Teda autokorelačné a parciálne autokorelačné funkcie danej rady sa v bode $k = 1$ vždy zhodujú.

Odhady r_{kk} autokorelačnej funkcie ϱ_{kk} sa obvykle počítajú rekurentne s využitím vzorca:

$$r_{11} = r_1,$$

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j}, \quad \text{pre } k > 1$$

kde

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j}, \quad \text{pre } j = 1, 2, \dots, k-1$$

POZNÁMKA. Pre konkrétnu analyzovanú časovú radu autokorelačnú funkciu ϱ_k často nepoznáme. Preto je nutné mať predstavu ako dobre ju zastupuje odhadovaná autokorelačná funkcia r_k . Je dôležité vedieť spočítať smerodatnú odchýlku odhadu r_k autokorelačnej funkcie ϱ_k .

Za týmto účelom sa využíva *Bartlettova aproximácia*.

Za predpokladu normality procesu platí, ak je $\varrho_k = 0$ pre $k > k_0$ potom

$$\sigma(r_k) = \sqrt{\text{var}(r_k)} \sim \sqrt{\frac{1}{n} \left(1 + 2 \sum_{j=1}^{k_0} r_j^2\right)}, \quad k > k_0$$

Pri rozhodovaní, či platí $\varrho_k = 0$, porovnávame $|r_k|$ s odhadom $2\sigma(r_k)$.

Ak platí $|r_k| > 2\sigma(r_k)$, potom zamietame hypotézu $\varrho_k = 0$ na predom stanovenej hladine významnosti 5 percent. Využíva sa asymptotická normalita odhadu r_k .

Nulovosť hodnôt parciálnej autokorelačnej funkcie ϱ_{kk} sa testuje na základe *Quenouilleovej aproximácie*.

Ak proces vyhovuje autoregresnému modelu rádu k_0 a $\varrho_{kk} = 0$ pre $k > k_0$ potom

$$\sigma(r_{kk}) \sim \sqrt{\frac{1}{n}}, \quad k > k_0$$

Hypotézu, že $\varrho_{kk} = 0$, teda zamietame, pokiaľ $|r_{kk}| > 2\sigma(r_{kk})$.

Lineárny proces

DEFINÍCIA. Biely šum ϵ_t je postupnosť nekorelovaných náhodných veličín s nulovou strednou hodnotou a konštantným rozptylom.

DEFINÍCIA. *Lineárnym procesom* nazývame radu tvaru

$$y_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} \dots \quad (1)$$

kde ϵ_t je biely šum s rozptylom σ_ϵ^2 a ψ_j sú parametre.

Lineárny proces sa dá ekvivalentne prepísať využitím *operátora spätného posunutia* B , ktorý je definovaný ako

$$By_t = y_{t-1},$$

pričom je možná niekoľko násobná aplikácia operátora

$$B^j y_t = y_{t-j}.$$

Z (1) je možné písať skrátene

$$y_t = \psi(B) \epsilon_t,$$

kde

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots = 1 + \sum_{j=1}^{\infty} \psi_j B^j.$$

Dá sa ukázať, že postačujúca podmienka pre existenciu lineárneho procesu má tvar

$$\psi(B) \text{ konverguje pro } |B| \leq 1.$$

Za predpokladu konverencie naviac platí, že (1) je stacionárny a $E(y_t) = 0$.

Za istých podmienok je možné (1) písať v tvare

$$y_t = \pi_1 y_{t-1} + \pi_2 y_{t-2} + \dots + \epsilon_t. \quad (2)$$

S využitím operátora spätného posunutia B sa dá ekvivalentne písať

$$\pi(B)y_t = \epsilon_t,$$

kde

$$\pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \dots = 1 - \sum_{j=1}^{\infty} \pi_j B^j.$$

Pokiaľ lineárny proces umožňuje zápis v tvare (2), potom sa nazýva *invertibilný*. Postačujúca podmienka pre invertibilitu lineárneho procesu má tvar

$$\pi(B) \text{ konverguje pro } |B| \leq 1.$$

Vzťah medzi parametrami ψ_j a π_j je

$$\psi(B)\pi(B) = 1, \quad \text{dá sa tiež písať } \pi(B) = \psi^{-1}(B).$$

POZNÁMKA. V praxi majú oveľa väčší význam špeciálne prípady lineárnych procesov. Sú to modely MA, AR a ARMA.

Modely MA(q)

DEFINÍCIA. Proces klzavých súčtov rádu q značený ako MA(q) má tvar

$$y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$$

alebo tiež $y_t = \theta(B)\epsilon_t$, kde $\theta(B) = 1 + \sum_{j=1}^q \theta_j B^j$ je tzv. *operátor klzavých súčtov*. Pritom ϵ_t je biely šum a θ_j sú parametre.

Proces MA(q) je stacionárny pre ľubovoľnú voľbu jeho parametrov. Stredná hodnota je nulová a pre rozptyl σ_y^2 platí

$$\sigma_y^2 = \gamma_0 = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_\epsilon^2,$$

jeho autokorelačná funkcia ϱ_k má tvar

$$\varrho_k = \frac{\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, \quad k = 1, \dots, q$$

$$\varrho_k = 0, \quad k > q.$$

Platí teda, že pre autokorelačnú funkciu procesu MA(q) je hľadaný bod k_0 rovný q . Pre parciálne autokorelačné funkcie procesu MA(q) neexistuje hľadaný bod k_0 , no sú obmedzené geometricky klesajúcou postupnosťou alebo sínusovou krivkou s geometricky klesajúcou amplitúdou.

Proces MA(q) je invertibilný ak sú všetky korene polynómu $\theta(B)$ vonku jednotkového kruhu.

Proces MA(1)

Tento proces má tvar $y_t = \epsilon_t + \theta_1 \epsilon_{t-1}$.

POZNÁMKA. Podmienka invertibility: $|\theta_1| < 1$.

Autokorelačná funkcia procesu MA(1) má tvar

$$\varrho_1 = \frac{\theta_1}{1 + \theta_1^2},$$

$$\varrho_k = 0, \quad k > 1 \quad (\text{identifikačný bod } k_0 = 1)$$

Proces MA(2)

Tento proces má tvar $y_t = \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2}$.

POZNÁMKA. Podmienka invertibility:

$$\theta_2 + \theta_1 > -1, \quad \theta_2 - \theta_1 > -1,$$

$$-1 < \theta_2 < 1.$$

Autokorelačná funkcia procesu MA(2) má tvar

$$\varrho_1 = \frac{\theta_1(1 + \theta_2)}{1 + \theta_1^2 + \theta_2^2},$$

$$\varrho_2 = \frac{\theta_2}{1 + \theta_1^2 + \theta_2^2},$$

$$\varrho_k = 0, \quad k > 2 \quad (\text{identifikačný bod } k_0 = 2)$$

POZNÁMKA. Máme daný ľubovoľný (teda aj neinvertibilný) proces MA(q). Potom platí

$$|\varrho_k| \leq \lim_{x \rightarrow q^-} \cos\left(\frac{\pi}{\lceil \frac{x+1}{k} \rceil + 2}\right)$$

lomené zátvorky značia celú časť čísla.

Špeciálne u modelu MA(2) platí $|\varrho_1| \leq \frac{1}{\sqrt{2}}, |\varrho_2| \leq \frac{1}{2}$.

Autoregresný proces AR(p)

DEFINÍCIA. Autoregresný proces rádu p značený ako AR(p) je definovaný ako

$$y_t = \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \epsilon_t \quad (3)$$

alebo ekvivalentne pomocou symboliky operátora spätného posunutia B ako

$$\varphi(B)y_t = \epsilon_t, \quad \text{kde } \varphi(B) = 1 - \sum_{j=1}^p \varphi_j B^j$$

je tzv. autoregresný operátor.

Proces AR(p) je invertibilný. Stacionárny je v prípade, že všetky korene polynómu $\varphi(B)$ ležia vonku jednotkového kruhu.

POZNÁMKA. Za platnosti podmienky stacionarity je možné prepísať do tvaru

$$y_t = \epsilon_t + \varphi_1 \epsilon_{t-1} + \varphi_1^2 \epsilon_{t-2} + \dots$$

Stredná hodnota stacionárneho procesu AR(p) je nulová. Vzťah (3) postupne násobíme veličinami y_{t-k} pre $k > 0$ a prejdeme k stredným hodnotám. Vieme, že $E(y_{t-k}\epsilon_t) = 0$ pre $k > 0$. Rovnosti napokon vydelíme γ_0 a dostaneme

$$\varrho_k = \varphi_1 \varrho_{k-1} + \varphi_2 \varrho_{k-2} + \dots + \varphi_p \varrho_{k-p}, \quad k > 0.$$

Na výpočet parametrov modelu AR(p) pomocou hodnôt autokorelačnej funkcie sa používa *Yule-Walkerova sústava rovníc* pre $k = 1, \dots, p$

$$\varrho_1 = \varphi_1 + \varphi_2 \varrho_1 + \dots + \varphi_p \varrho_{p-1},$$

$$\varrho_2 = \varphi_1 \varrho_1 + \varphi_2 + \dots + \varphi_p \varrho_{p-2},$$

$$\vdots$$

$$\varrho_p = \varphi_1 \varrho_{p-1} + \varphi_2 \varrho_{p-2} + \dots + \varphi_p.$$

Riešením sústavy budú parametre $\varphi_1, \dots, \varphi_p$ vyjadrené pomocou autokorelačných hodnôt $\varrho_1, \dots, \varrho_p$. Rozptyl procesu AR(p) je rovný

$$\sigma_y^2 = \gamma_0 = \frac{\sigma_\epsilon^2}{1 - \varphi_1 \varrho_1 - \dots - \varphi_p \varrho_p}.$$

Platí, že pre parciálnu autokorelačnú funkciu procesu AR(p) je $k_0 = p$, teda $\varrho_{kk} = 0$ pre $k > p$.

Proces AR(1)

Tento proces má tvar $y_t = \varphi_1 y_{t-1} + \epsilon_t$.

POZNÁMKA. Podmienka staconarity: $|\varphi_1| < 1$.

Autokorelačná funkcia procesu AR(1) má tvar

$$\varrho_k = \varphi_1^k, \quad k \geq 0 \quad \text{špeciálne je } \varphi_1 = \varrho_1.$$

Parciálna autokorelačná funkcia procesu AR(p) má identifikačný bod $k_0 = 1$ a teda $\varrho_{11} = \varphi_1$ a $\varrho_{kk} = 0$ pre $k > 1$.

Proces AR(2)

Tento proces má tvar $y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \epsilon_t$.

POZNÁMKA. Podmienka stacionarity procesu je tvaru

$$\varphi_2 + \varphi_1 < 1, \quad \varphi_2 - \varphi_1 < 1, \quad -1 < \varphi_2 < 1.$$

Yule-Walkerove rovnice tvaru

$$\varrho_1 = \varphi_1 + \varphi_2 \varrho_1,$$

$$\varrho_2 = \varphi_1 \varrho_1 + \varphi_2$$

nám dávajú riešenie

$$\varphi_1 = \frac{\varrho_1(1 - \varrho_2)}{1 - \varrho_1^2}, \quad \varphi_2 = \frac{\varrho_2 - \varrho_1^2}{1 - \varrho_1^2}.$$

Z čoho môžeme pre stacionárny proces AR(2) odvodiť vzťahy

$$-1 < \varrho_2 < 1, \quad \varrho_1^2 < \frac{\varrho_2 + 1}{2}.$$

Pre identifikačný bod procesu platí $k_0 = 2$.

Zmiešaný proces ARMA(p,q)

DEFINÍCIA. Zmiešaný proces rádu p a q, tiež značený ako ARMA(p,q) definujeme ako

$$y_t = \varphi_1 y_{t-1} + \dots + \varphi_p y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q}$$

alebo pomocou symboliky operátora spätného posunutia

$$\varphi(B)y_t = \theta(B)\epsilon_t.$$

POZNÁMKA. Stacionarita a invertibilita procesu ARMA(p,q) je daná podmienkou stacionarity procesu AR(p) a podmienkou invertibility procesu MA(q). Stredná hodnota procesu ARMA(p,q) je nulová.

Pre autokorelačnú funkciu ϱ_k platí

$$\varrho_k = \varphi_1 \varrho_{k-1} + \varphi_2 \varrho_{k-2} + \dots + \varphi_p \varrho_{k-p}, \quad k > p. \quad (4)$$

Dá sa ukázať, že autokorelačná funkcia je lineárnou kombináciou klesajúcich geometrických postupností a sínusoid s geometricky klesajúcou amplitúdou, rovnako ako v prípade procesu AR(p), ale pre $k \geq \max(0, p - q + 1)$. Parciálna autokorelačná funkcia ϱ_{kk} procesu ARMA(p,q) je rovnaká ako v prípade procesu MA(q), len pre $k \geq \max(1, p - q + 1)$. Špeciálne teda neexistuje identifikačný bod k_0 .

Proces ARMA(1,1)

Tento proces má tvar $y_t = \varphi_1 y_{t-1} + \epsilon_t + \theta_1 \epsilon_{t-1}$.

POZNÁMKA. Podmienka stacionarity má tvar $|\varphi_1| < 1$, podmienka invertibility $|\theta_1| < 1$.

Sústava (4) má pre proces ARMA(p,q) tvar

$$\varrho_k = \varphi_1 \varrho_{k-1}, \quad k > 1.$$

Pre rekurentný výpočet ϱ_k pomocou (4) je potrebné poznať ϱ_1 . Tú určíme pomocou γ_1 a γ_0 . Rovnicu procesu ARMA(1,1) vynásobíme veličinami y_t , y_{t-1} a prejdeme ku stredným hodnotám.

Po úpravách dostaneme

$$\sigma_y^2 = \gamma_0 = \frac{1 + \theta_1^2 + 2\varphi_1\theta_1}{1 - \varphi_1^2} \sigma_\epsilon^2,$$

$$\gamma_1 = \frac{(1 + \varphi_1\theta_1)(\varphi_1 + \theta_1)}{1 - \varphi_1^2} \sigma_\epsilon^2.$$

Odtiaľ

$$\varrho_1 = \frac{(1 + \varphi_1\theta_1)(\varphi_1 + \theta_1)}{1 + \theta_1^2 + 2\varphi_1\theta_1}.$$

POZNÁMKA. Parciálna autokorelačná funkcia procesu ARMA(1,1) je rovnako ako v prípade MA(1) obmedzená geometricky klesajúcou postupnosťou.

POZNÁMKA. V praxi sa obvykle vystačí s procesmi, pre ktoré je $p + q \leq 2$.

1.3. Tri fázy tvorby modelu

Pred fázou identifikácie je vhodné previesť prípravné operácie: zistiť stacionaritu a nulovú strednú hodnotu.

- (1) **Identifikácia** Identifikácia je prvou fázou výstavby modelu a jej úlohou je rozhodnúť aký typ modelu vybrať a explicitne určiť rád modelu.

Rozhodovacie schéma

	AR(p)	MA(q)	ARMA(p,q)
ρ_k	neexistuje k_0 , ρ_k v tvare U	$k_0 = q$	neexistuje k_0 , ρ_k v tvare U po prvých q-p hodnotách
ρ_{kk}	$k_0 = p$	neexistuje k_0 , ρ_{kk} obmedzená krivkou v tvare U	neexistuje k_0 , ρ_{kk} obmedzená krivkou v tvare U po prvých p-q hodnotách

Symbol U značí krivku v tvare lineárnej kombinácie klesajúcich geometrických postupností a sínusoid s geometricky klesajúcou amplitúdou. Pre k_0 platí, že pre $|k| > k_0$ je $\rho_k = 0$, resp. $\rho_{kk} = 0$.

Pokiaľ sa nepodari zvoliť žiadny AR ani MA model, je možné zvoliť nejaký z modelov ARMA a pre tento model previesť odhad parametrov a verifikáciu modelu. Ak sú zvolené rády príliš vysoké, prejaví sa táto skutočnosť na hodnotách odhadov parametrov, ktoré budú štatisticky nevýznamné. Príliš nízke rády sa odhalia vo fáze verifikácie modelu.

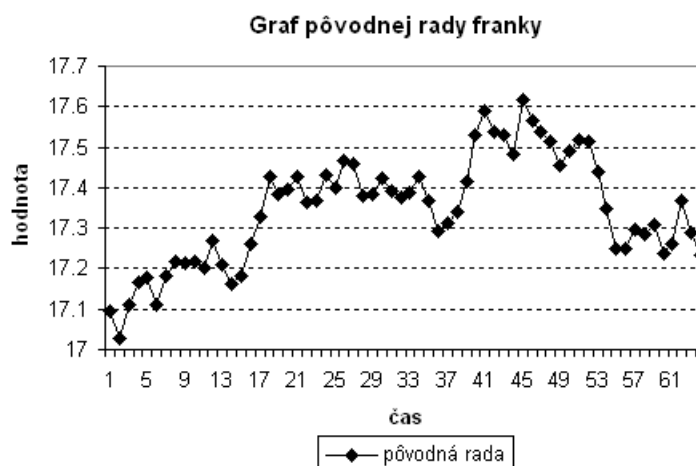
Software NCSS tiež môže sám navrhnúť optimálny model na základe minimalizácie reziduálneho súčtu štvorcov.

- (2) **Odhad parametrov** V súčasnej dobe sa používa mnoho odhadových procedúr. V NCSS je implementovaná iteračná procedúra založená na metóde maximálnej vierohodnosti vychádzajúca z predpokladu normality bieleho šumu.
- (3) **Verifikácia modelu** Overovanie konštruovaného modelu spočíva v analýze rezíduí, teda odhadnutých hodnôt bieleho šumu. Posudzuje sa ich nekorelovanosť a normalita.

1.4. Analýza dát

Praktická časť bakalárskej práce je založená na nájdení vhodného modelu, ktorý by spoľahlivo aproximoval budúci vývoj (tj. v našom prípade 5 pozorovaní) nejakej nami vybranej rady. Spracovávame radu 64 hodnôt kurzu švajčiarskeho franku k českej korune v období od 2.1.2007 do 30.3.2007. Údaje tvoria dňové pozorovania a boli získané z webových stránok Českej národnej banky.

Predpovede do budúcnosti boli konštruované v skrátenej rade a porovnané so skutočným vývojom. Najprv uvedieme výsledky Boxovej-Jenkinsovej metodológie, potom dvojitého exponenciálneho vyrovnávania.



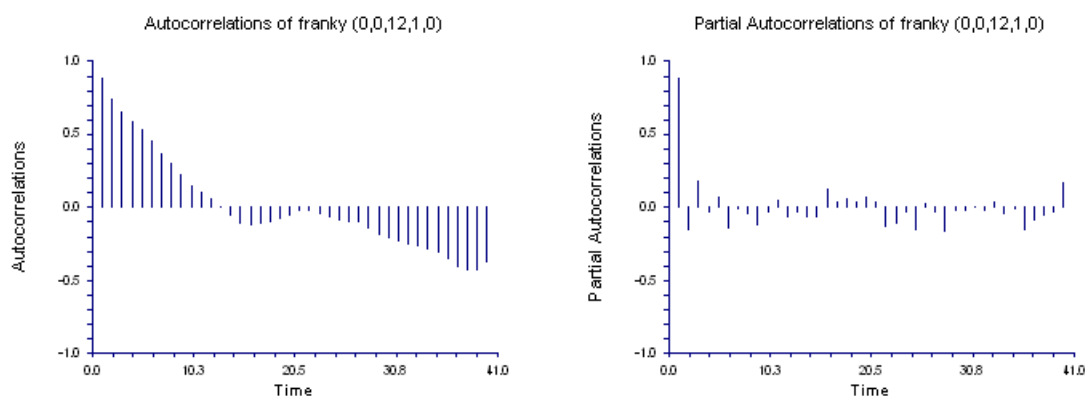
OBR. 1. Graf pôvodnej rady frankov

Boxova-Jenkinsova metodológia

V programe NCSS bola zvolená transformácia, ktorá odpočíta priemer (tj. Remove Mean).

Proces hľadania modelu

Pre základnú orientáciu v probléme voľby modelu, je potrebné poznať hodnoty autokorelačnej (ACF) a parciálnej autokorelačnej (PACF) funkcie a vedieť ich interpretovať.



OBR. 2. Autokorelačná a parciálna autokorelačná funkcia

Na obrázku je vidieť, že PACF je useknutá v jednej, pre ostatné hodnoty PACF neplatí podmienka uvedená na výstupe z NCSS

Significant if $|\text{Correlation}| > 0.250000$,

teda nezamietame nulovosť parciálnych korelácií ρ_{kk} pre $k \geq 2$ v teste založenom na použití Quenouillovej aproximácie, kde $n = 64$ a $2\sigma(r_{kk}) = 0.25$

Vzhľadom k tvaru ACF dostávame výstup typický pre model AR(1).

Prehľad číselných hodnôt autokorelácií a parciálnych autokorelácií.

Lag	Correlation	Lag	Correlation	Lag	Correlation	Lag	Correlation
1	0,877615	9	0,225598	17	-0,108385	25	-0,078955
2	0,735557	10	0,148800	18	-0,095448	26	-0,092675
3	0,651160	11	0,101803	19	-0,076819	27	-0,099254
4	0,582681	12	0,056287	20	-0,052248	28	-0,140835
5	0,528258	13	0,006547	21	-0,020004	29	-0,180888
6	0,451355	14	-0,047500	22	-0,015810	30	-0,205287
7	0,371140	15	-0,106577	23	-0,042861		
8	0,304956	16	-0,120179	24	-0,058019		

TABUĽKA 1. Odhadnuté autokorelácie

Lag	Correlation	Lag	Correlation	Lag	Correlation	Lag	Correlation
1	0,877615	9	-0,113430	17	0,033678	25	-0,145047
2	-0,150799	10	-0,034402	18	0,058190	26	0,030215
3	0,180889	11	0,051763	19	0,035390	27	-0,024814
4	-0,028434	12	-0,061074	20	0,072411	28	-0,161166
5	0,066433	13	-0,031018	21	0,037566	29	-0,014169
6	-0,141526	14	-0,057699	22	-0,126990	30	-0,021338
7	-0,002514	15	-0,061661	23	-0,104464		
8	-0,037014	16	0,123508	24	-0,025238		

TABUĽKA 2. Odhadnuté parciálne autokorelácie

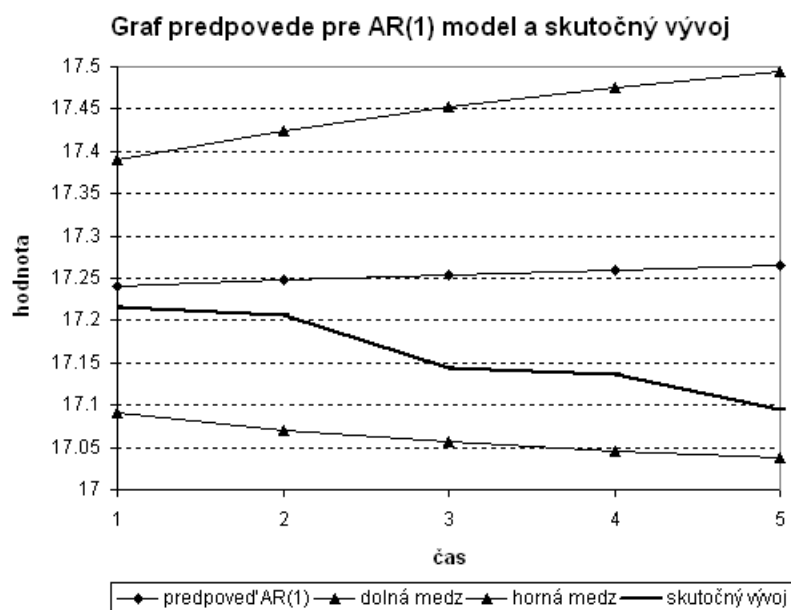
Model AR(1)

Model AR(1) je definovaný rovnicou $y_t = \varphi_1 y_{t-1} + \epsilon_t$, parameter φ_1 odhadujeme hodnotou $\hat{\varphi}_1$, ϵ_t je náhodná zložka.

Parameter Name	Parameter Estimate	Standard Error	T-value	Prob Level
AR(1)	0,9334589	0,041299	22,6025	0,000000

TABUĽKA 3. Výstup z programu NCSS

Odhad parametru pre model je $\hat{\varphi}_1 = 0.933$, je splnená stacionarita procesu. Priemer štvorcov rezíduí je rovný 0.003105765. V modeli je testovaná hypotéza $H_0 : \varphi_1 = 0$ na hladine 5 percent. Probability level (v našom prípade hodnota 0.000000) je najmenšia hladina, na ktorej zamietame hypotézu H_0 .



OBR. 3. Model AR(1) a skutočný vývoj

Do budúcnosti je predpokladaný mierny nárast. Okolo bodových predpovedí sú konštruované 95 percentné intervaly spoľahlivosti vytvárajúce pás spoľahlivosti (v grafe označený ako dolná a horná medz).

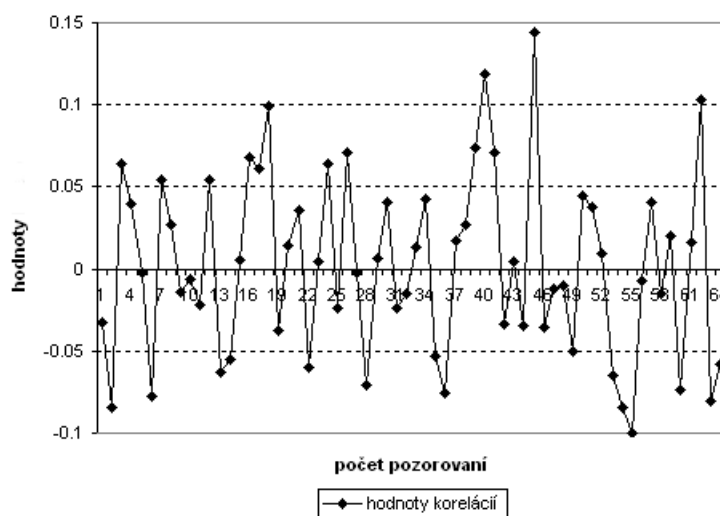
Verifikácia modelu

Hodnoty ACF rezíduí môžeme považovať za nulové, nespĺňajú podmienku,

$$\text{Significant if } |\text{Correlation}| > 0.250000.$$

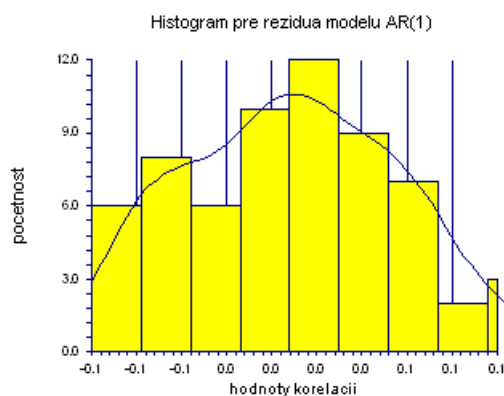
Lag	Correlation	Lag	Correlation	Lag	Correlation	Lag	Correlation
1	0.065639	13	-0.048497	25	-0.027818	37	-0.154340
2	-0.186719	14	0.007812	26	-0.084328	38	-0.022370
3	-0.085020	15	-0.205973	27	0.096916	39	-0.158733
4	-0.075379	16	-0.050811	28	-0.014472	40	-0.007673
5	0.088900	17	0.106394	29	-0.004270	41	0.129292
6	0.078550	18	-0.072846	30	-0.014793	42	0.039675
7	-0.031087	19	-0.002157	31	-0.079035	43	-0.104840
8	0.040528	20	-0.085147	32	-0.020664	44	0.033565
9	-0.012772	21	0.104168	33	0.104972	45	0.005798
10	-0.133877	22	0.172216	34	0.064157	46	-0.038553
11	0.099278	23	0.022005	35	0.033345	47	0.018902
12	-0.015577	24	-0.037139	36	0.039461	48	-0.071695

TABUĽKA 4. Odhadnuté autokorelácie rezíduí



OBR. 4. Hodnoty rezíduí v modeli AR(1)

Rezídua kmitajú okolo nulovej úrovne, ich strednú hodnotu môžeme považovať za nulovú.



OBR. 5. Histogram rezíduí AR(1)

Histogram pre rezídua vykazuje približne priebeh hustoty normálneho rozdelenia so strednou hodnotou rovnou nule. Pre potvrdenie alebo zamietnutie normality rezíduí bolo spravených niekoľko testov, žiadny z nich hypotézu normality na 5 percentnej hladine nezamieta.

Test Name	Test Value	Prob Level	10 percent Critical Value	Decision 5 percent
<i>Shapiro – Wilk</i>	0.9819139	0.4699838		<i>Can't reject normality</i>
<i>Anderson – Darling</i>	0.2542238	0.7307234		<i>Can't reject normality</i>
<i>Martinez – Iglewicz</i>	0.9597874		1.075744	<i>Can't reject normality</i>
<i>Kolmogorov – Smirnov</i>	0.0592516		0.101	<i>Can't reject normality</i>
<i>D'Agostino Skewness</i>	0.9142016	0.3606109	1.645	<i>Can't reject normality</i>
<i>D'Agostino Kurtosis</i>	−0.6877	0.491636	1.645	<i>Can't reject normality</i>
<i>D'Agostino Omnibus</i>	1.3087	0.519778	4.605	<i>Can't reject normality</i>

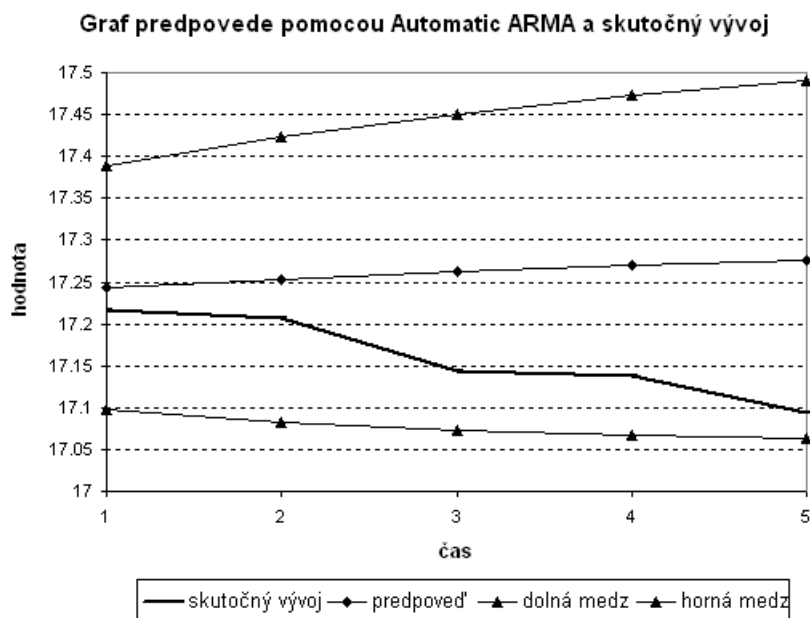
TABUĽKA 5. Testy normality

Program NCSS ďalej ponúka procedúru automatického vyhľadávania optimálneho modelu pre dané dáta. Tá je založená na prechádzaní rôznych ARMA modelov a na minimalizácii súčtu štvorcov rezíduí.

Model Automatic ARMA

V procedúre Automatic ARMA sa zadala transformácia odpočítania priemeru, ostatné parametre automatického hľadania boli automaticky generované programom.

Za najvhodnejší model pre aproximáciu bol iteračne zvolený model AR(1). MSE je rovné 0.00304216. Odhad parametru $\hat{\varphi}_1 = 0.9078866$. V budúcnosti sa predpokladá mierny rast.

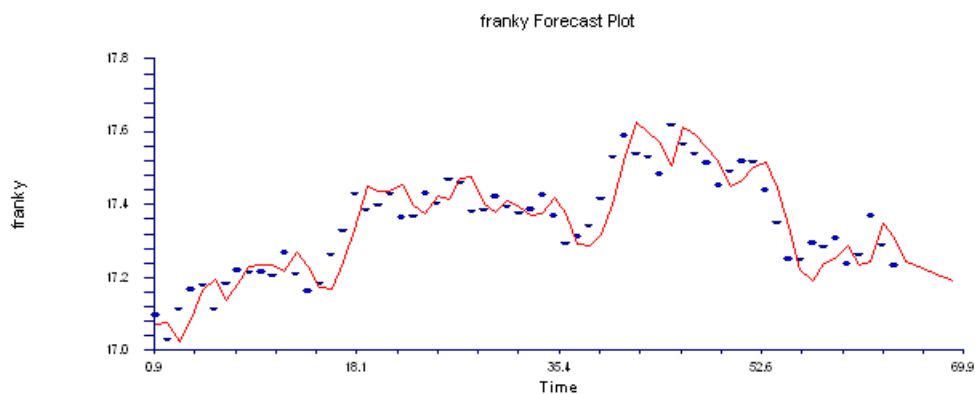


OBR. 6. Model Automatic ARMA a skutočný vývoj

Výsledky verifikácie v modeli Automatic ARMA sú totožné s modelom AR(1) navrhnutým na základe štúdia korelačných funkcií časovej rady. Mierna odlišnosť odhadu $\hat{\varphi}_1$ v oboch prístupoch je pravdepodobne spôsobená rozdielnosťami v implementovaní odhadovej procedúry v príslušných moduloch NCSS. Manuály programu bohužiaľ neponúkajú podrobnejšie informácie o výpočte odhadov.

Dvojité exponenciálne vyrovňavanie

Teraz uvádzame výsledky spracovania rovnakej časovej rady metódou dvojitého exponenciálneho vyrovňavania v programe NCSS. V obrázku vidíme body reprezentujúce pôvodné dáta a čiaru vyrovnanej rady s naviazanými predpoveďami 5 budúcich hodnôt.



OBR. 7. Dvojité exponenciálne vyrovňavanie

MSE je 0.004047862, čo je viac ako v prípade modelov AR(1) a Automatic ARMA, no táto hodnota je daná spôsobom tvorby modelu. Graf vyrovnanej rady dvojitého exponenciálneho vyrovňavania je totižto posunutý oproti pôvodnej rade o určitú konštantu doprava v smere x-ovej osi.

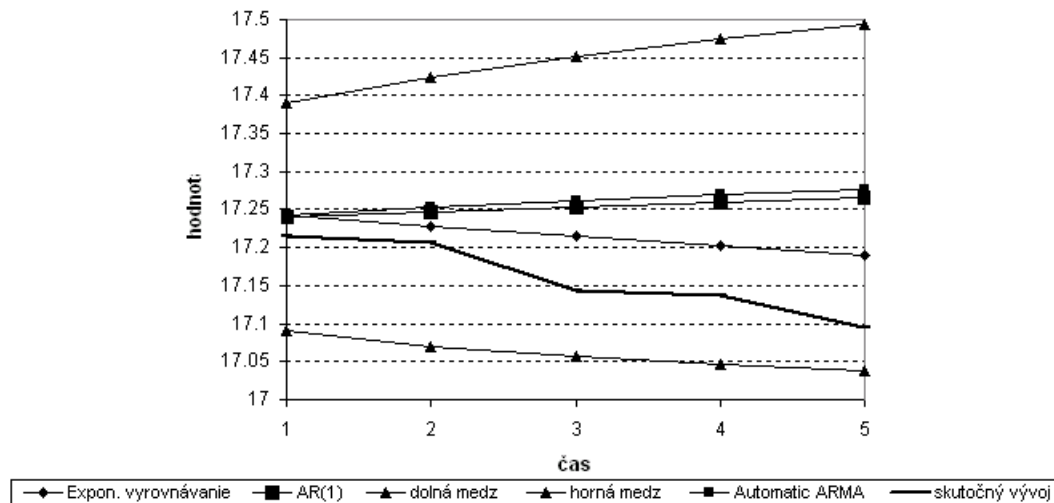
V nasledujúcej tabuľke sú hodnoty parametrov α , a_0 a b_0 . Sú to parametre priamky, ktorá predpovedá budúci vývoj.

<i>Alpha</i>	0.4638187
<i>Intercept(A)</i>	18.07877
<i>Slope(B)</i>	-0.0128777

TABUĽKA 6. Parametre dvojitého exponenciálneho vyrovňavania

Záverom porovnajme predpovedanie do budúcnosti vo všetkých aplikovaných modeloch. Pás spoľahlivosti sa týka predpovedí z modelu AR(1).

Graf porovnávajúci model Automatic ARMA, AR(1) a Exponenciálne vyrovňovanie



OBR. 8. Porovnanie predpovedí

Záver

Napriek dostatočnému počtu pozorovaní v časovej rade a sofistikovanejšej tvorbe predpovedí a voľbe parametrov, Boxov-Jenkinsov model neodhadol budúci priebeh vývoja meny lepšie ako model dvojitého exponenciálneho vyrovňovania. V skutočnosti mal kurz v nasledujúcich piatich pozorovaniach klesajúci trend, najlepšou aproximáciou je v tomto prípade dvojité exponenciálne vyrovňovanie. V modeli AR(1) je skutočný vývoj pokrytý pásom spoľahlivosti.

KAPITOLA 2

Mnohorozmerné rady

OZNAČENIE. V práci budú uvažované stĺpcové vektory a budú značené tučným písmom \mathbf{X} , transpozície budú \mathbf{X}^T , stredná hodnota bude $E\mathbf{X}$, rozptylová matica bude $\text{var}\mathbf{X}$ a kovariančná matica náhodných vektorov \mathbf{X} a \mathbf{Y} sa bude značiť ako $\text{cov}(\mathbf{X}, \mathbf{Y})$

OZNAČENIE. Keďže sa celá kapitola zaoberá mnohorozmernými prípadmi, v ďalšom texte budeme slovo m-rozmerný vynechávať.

2.1. Základné vlastnosti mnohorozmerného náhodného procesu

DEFINÍCIA. Nech $T \subset \mathbb{R}$ a pre všetky $t \in T$ je \mathbf{X}_t reálny náhodný vektor. Potom sa systém $\{\mathbf{X}_t, t \in T\}$ nazýva *náhodný proces*.

Špeciálne ak je T podmnožina celých čísel, nazýva sa $\{\mathbf{X}_t, t \in T\}$ *náhodná postupnosť* alebo tiež *časová rada*.

Ak je z kontextu jasné o akú množinu T sa jedná, budeme písať len \mathbf{X}_t

DEFINÍCIA. Nech $\{\mathbf{X}_t, t \in T\}$ je náhodný proces taký, že pre každé $t \in T$ existuje stredná hodnota $E\mathbf{X}_t$. Potom sa vektorová funkcia $\mathbf{e}_t = E\mathbf{X}_t$ definovaná na T nazýva *stredná hodnota* náhodného procesu $\{\mathbf{X}_t\}$.

Náhodný proces s nulovou strednou hodnotou sa nazýva *centrovaný náhodný proces*.

DEFINÍCIA. Nech $\{\mathbf{X}_t, t \in T\}$ je náhodný proces so strednou hodnotou $E\mathbf{X}_t$, kde $\mathbf{X}_t = (X_{1,t}, X_{2,t}, \dots, X_{m,t})^T$, a platí $E(X_{k,t})^2 < \infty$ pre $t \in T$ a $k = 1, \dots, m$, potom sa maticová funkcia $\Gamma(s, t) = E(\mathbf{X}_s - E\mathbf{X}_s)(\mathbf{X}_t - E\mathbf{X}_t)^T$ definovaná na $T \times T$ nazýva *kovariančná maticová funkcia* náhodného procesu $\{\mathbf{X}_t\}$. Funkcia $\Gamma(t, t)$ definovaná na T sa nazýva *rozptylová maticová funkcia* náhodného procesu $\{\mathbf{X}_t\}$.

DEFINÍCIA. Ak je stredná hodnota nejakého náhodného procesu konštantná a kovariančná maticová funkcia $\Gamma(s, t)$ je len funkciou rozdielu argumentov $t - s$ a je konečná pre všetky $s, t \in T$, potom sa tento náhodný proces nazýva *slabo stacionárny*. Potom skrátene píšeme $\Gamma(k) = \Gamma(t, t+k)$ a $\Gamma(0)$ nazývame *rozptylová matica*. Naďalej budeme uvažovať len procesy s nulovou strednou hodnotou.

POZNÁMKA. Pre $t \in \mathbb{Z}$ platí $\Gamma(-k) = \Gamma(k)^T$. Preto maticu často počítame len pre $t \in \mathbb{N}_0$.

2.2. Mnohorozmerné ARMA modely

Definícia postupnosti ARMA

DEFINÍCIA. Nech $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ je náhodná postupnosť, $E\mathbf{Y}_t = \mathbf{0}$ pre všetky $t \in \mathbb{Z}$, kovariančná maticová funkcia $\Gamma(s, t) = \mathbf{0}$ je nulová matica pre všetky $s \neq t$, $s, t \in \mathbb{Z}$ a $\Gamma(t, t)$ je konečná, symetrická, pozitívne definitná matica a je rovná pre všetky $t \in \mathbb{Z}$. Potom $\{\mathbf{Y}_t\}$ nazývame *biely šum*.

POZNÁMKA. Biely šum je centrovaná slabo stacionárna náhodná postupnosť, ktorej kovariančná maticová funkcia $\Gamma(k)$ je nulová pre všetky $k \in \mathbb{Z} \setminus 0$.

Symbol \mathbf{Y}_t bude značiť m -rozmerný biely šum.

Rozptylová matica $\Gamma(0)$ bieleho šumu bude značená Σ .

POZNÁMKA. Rovnako ako v prípade jednorozmerných rád zavádzame operátor spätného posunutia \mathbf{B} , ktorý je definovaný predpisom

$$\mathbf{B}^j \mathbf{X}_t = \mathbf{X}_{t-j}, \quad j = 1, 2, \dots$$

DEFINÍCIA. Nech $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ je biely šum, Ψ_1, Ψ_2, \dots sú matice typu $m \times m$. Lineárnym procesom nazývame náhodnú postupnosť tvaru

$$\mathbf{X}_t = \mathbf{Y}_t + \Psi_1 \mathbf{Y}_{t-1} + \Psi_2 \mathbf{Y}_{t-2} + \dots$$

Pomocou operátora spätného posunutia môžeme písať $\mathbf{X}_t = \Psi(\mathbf{B})\mathbf{Y}_t$, kde $\Psi(\mathbf{B}) = \sum_{s=0}^{\infty} \Psi_s \mathbf{B}^s$ a $\Psi_0 = I$.

Stacionaritu procesu zaručí podmienka $\sum_{s=0}^{\infty} \psi_{ij,s}^2 < \infty$ pre každé $i, j = 1, \dots, m$. Invertibilitou rozumieme možnosť zápisu procesu \mathbf{X}_t v tvare

$$\mathbf{X}_t = \Pi_1 \mathbf{X}_{t-1} + \Pi_2 \mathbf{X}_{t-2} + \dots + \mathbf{Y}_t$$

alebo $\Pi(\mathbf{B})\mathbf{X}_t = \mathbf{Y}_t$, kde $\Pi(\mathbf{B}) = I - \sum_{s=1}^{\infty} \Pi_s \mathbf{B}^s$ a $\Pi_0 = I$.

Invertibilitu procesu zaručí podmienka $\sum_{s=0}^{\infty} |\pi_{ij,s}| < \infty$ pre každé $i, j = 1, \dots, m$.

POZNÁMKA. Konvergenciou náhodných veličín a vektorov sa bude myslieť konvergencia podľa kvadratického stredu, teda napríklad symbol $\mathbf{X}_t = \sum_{s=0}^{\infty} \Psi_s \mathbf{Y}_{t-s}$ znamená

$$\lim_{q \rightarrow \infty} E[(\mathbf{X}_t - \sum_{s=0}^q \Psi_s \mathbf{Y}_{t-s})^T (\mathbf{X}_t - \sum_{s=0}^q \Psi_s \mathbf{Y}_{t-s})] = 0.$$

DEFINÍCIA. Nech $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ je biely šum, $q \in \mathbb{N}$, $\Theta_1, \dots, \Theta_q$ sú matice typu $m \times m$, pričom Θ_q je nenulová. Potom sa náhodná postupnosť $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ nazýva *postupnosť kľazavých súčtov rádu q* , označujeme MA(q) ak

$$\mathbf{X}_t = \mathbf{Y}_t + \Theta_1 \mathbf{Y}_{t-1} + \Theta_2 \mathbf{Y}_{t-2} + \dots + \Theta_q \mathbf{Y}_{t-q} \text{ pre všetky } t \in \mathbb{Z}$$

Pri označení $\Theta(\mathbf{B}) = \sum_{s=0}^q \Theta_s \mathbf{B}^s$, kde \mathbf{B} je operátor spätného posunutia a $\Theta_0 = I$ je jednotková matica, môžeme ekvivalentne písať

$$\mathbf{X}_t = \Theta(\mathbf{B})\mathbf{Y}_t.$$

DEFINÍCIA. Nech $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ je biely šum, $p \in \mathbb{N}$, $\Phi_1, \Phi_2, \dots, \Phi_p$ sú matice typu $m \times m$, pričom Φ_p je nenulová. Potom sa náhodná postupnosť $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ nazýva *autoregresná postupnosť rádu p* , značíme AR(p), ak

$$\mathbf{X}_t = \mathbf{Y}_t + \Phi_1 \mathbf{X}_{t-1} + \Phi_2 \mathbf{X}_{t-2} + \dots + \Phi_p \mathbf{X}_{t-p} \text{ pre všetky } t \in \mathbb{Z}$$

Ak opäť označíme $\Phi(\mathbf{B}) = \Phi_0 - \sum_{s=1}^p \Phi_s \mathbf{B}^s$, kde $\Phi_0 = I$ je jednotková matica, môžeme ekvivalentne písať

$$\Phi(\mathbf{B})\mathbf{X}_t = \mathbf{Y}_t.$$

OZNAČENIE. Matice Θ_0 a Φ_0 budeme stotožňovať s jednotkovými maticami.

DEFINÍCIA. Nech $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ je biely šum, $p, q, \Theta_0, \Theta_1, \dots, \Theta_q, \Phi_0, \Phi_1, \dots, \Phi_p, \Theta(\mathbf{B})$ a $\Phi(\mathbf{B})$ sú ako v predchádzajúcich definíciách, potom sa náhodná postupnosť $\{\mathbf{X}_t, t \in \mathbb{Z}\}$ nazýva *zmiešaný model autoregresie a klzavých súčtov rádu p a q*. Označujeme ako ARMA(p,q) a môžeme písať

$$\Phi(\mathbf{B})\mathbf{X}_t = \Theta(\mathbf{B})\mathbf{Y}_t.$$

Stacionarita a invertibilita postupností ARMA

Teória ARMA modelov sa vzťahuje na triedu postupností s vlastnosťou stacionarity a invertibility, ktoré sú dôležité pri identifikácii modelu. Vo väčšine prípadov analýzy časových rád je možné transformovať nestacionárnu radu na stacionárnu.

Postupnosť klzavých súčtov je stacionárna.

Invertibilná je, pokiaľ korene polynómu $\det(\Theta(\mathbf{B}))$ ležia vonku jednotkového kruhu. Autoregresná postupnosť je invertibilná.

Stacionárna je, pokiaľ korene polynómu $\det(\Phi(\mathbf{B}))$ ležia vonku jednotkového kruhu. Stacionarita a invertibilita postupnosti ARMA(p,q) je zaručená, pokiaľ korene polynómu $\det(\Phi(\mathbf{B}))$ resp. $\det(\Theta(\mathbf{B}))$ ležia vonku jednotkového kruhu.

Stacionarizácia časovej rady

Stacionárna ARMA postupnosť sa vyznačuje tým, že jej stredná hodnota je nulová a jej kovariančná maticová funkcia sa nemení v čase. Ak časová rada vykazuje buď nenulovú strednú hodnotu alebo nekonštantnú kovariančnú maticovú funkciu, využívajú sa k stacionarizácii rady metódy analogické metódam používaným v jednorozmernom prípade, keďže mnohorozmernú časovú radu je možné transformovať po zložkách. Najčastejšie používanou transformáciou je prechod k rade diferencií.

2.3. Identifikácia ARMA modelu

Rovnako ako v jednorozmernom prípade, aj v mnohorozmerných modeloch sa uprednostňujú modely nižších rádo.

Pred vlastnou identifikáciou modelu by mala byť pozorovaná rada transformovaná na radu stacionárnu a centrovanú.

OZNAČENIE. Pojmom časová rada sa bude rozumieť pozorovaná rada. Hodnota v čase t sa bude značiť \mathbf{x}_t . V teoretickom popise sa bude používať pojem náhodná postupnosť a náhodné vektory budeme značiť \mathbf{X}_t .

Zložky týchto jednotlivých vektorov budeme značiť $x_{i,t}$ resp. $X_{i,t}$ pre $i = 1, 2, \dots, m$. Rozsah výberu \mathbf{x}_t budeme označovať n , tzn. predpokladáme, že poznáme \mathbf{x}_t pre $t = 1, 2, \dots, n$.

Určenie rádu modelu

Určovanie rádu p a q modelu ARMA(p,q) je založené na skúmaní výberovej korelačnej maticovej funkcie, výberovej parciálnej autoregresnej maticovej funkcie a výberovej parciálnej korelačnej maticovej funkcie. Obvykle stačí použiť len niekoľko prvých hodnôt týchto funkcií, nemalo by sa ich používať viac ako $\frac{n}{4}$. Odhady príslušných charakteristík môžu byť skreslené, preto v prípade nejasej voľby medzi niekoľkými modelmi je vhodné pokračovať so všetkými. Pre ten najlepší model sa rozhodneme až vo fáze verifikácie modelu.

2.3.1. Kovariančná a korelačná maticová funkcia.

Ak chceme identifikovať pre pozorovanú časovú radu model $MA(q)$, je vhodné použiť kovariančnú a korelačnú maticovú funkciu.

Dá sa ukázať, že kovariančná maticová funkcia $\Gamma(k)$ postupnosti $MA(q)$ je nulová pre všetky $|k| > q$.

DEFINÍCIA. Nech $\{\mathbf{X}_t\}$ je slabo stacionárna náhodná postupnosť s kovariančnou maticovou funkciou $\Gamma(k) = [\gamma_{ij}(k)]_{i,j=1,2,\dots,m}$. Označme Δ diagonálnu maticu

$$\Delta = \text{diag}\left\{\frac{1}{\sqrt{\gamma_{11}(0)}}, \frac{1}{\sqrt{\gamma_{22}(0)}}, \dots, \frac{1}{\sqrt{\gamma_{mm}(0)}}\right\}.$$

Potom pre každé $k \in \mathbb{Z}$ definujeme *korelačnú maticu* $\rho(k)$ ako

$$\rho(k) = \Delta \Gamma(k) \Delta.$$

POZNÁMKA. Matica $\rho(k)$ má na mieste i, j prvok

$$\rho_{ij}(k) = \frac{\gamma_{ij}(k)}{\sqrt{\gamma_{ii}(0)\gamma_{jj}(0)}},$$

ktorý je korelačným koeficientom i -tej zložky vektoru \mathbf{X}_t a j -tej zložky vektoru \mathbf{X}_{t+k} . Teda korelačná matica je nulová, práve vtedy keď je nulová kovariančná matica.

DEFINÍCIA. Nech časová rada $\{\mathbf{x}_t, t = 1, 2, \dots, n\}, n \in \mathbb{N}$ je realizáciou náhodnej postupnosti $\mathbf{X}_t, t \in \mathbb{N}$. Potom sa matica s prvkami

$$\hat{\rho}_{ij}(k) = \frac{\sum_{t=1}^{n-k} (x_{t,i} - \bar{x}_i)(x_{t+k,j} - \bar{x}_j)}{(\sum_{t=1}^n (x_{t,i} - \bar{x}_i)^2 \sum_{t=1}^n (x_{t,j} - \bar{x}_j)^2)^{\frac{1}{2}}}$$

nazýva *výberová korelačná matica*.

POZNÁMKA. Za odhad strednej hodnoty pre jednotlivé zložky časovej rady budeme považovať $\bar{x}_i = \frac{1}{n} \sum_{t=1}^n x_{t,i}, i = 1, 2, \dots, m$.

Vlastnosti výberovej korelačnej matice

Pokiaľ pre prvky korelačnej maticovej funkcie platí $\rho_{ij}(k) = 0$ pre $|k| > q$ pre nejaké q , potom pre $k > q$ a pre veľké n má $\hat{\rho}_{ij}(k)$ asymptoticky normálne rozdelenie s nulovou strednou hodnotou a rozptylom

$$\sigma_{ij}(k)^2 = \frac{1}{n-k} \left(1 + 2 \sum_{s=1}^k \rho_{ii}(s) \rho_{jj}(s)\right).$$

Naviac pre $k > q$ má štatistika

$$R(k) = \sum_{i=1}^m \sum_{j=1}^m \left(\frac{\hat{\rho}_{ij}(k)}{\sigma_{ij}(k)}\right)^2$$

asymptoticky chí-kvadrát rozdelenie o m^2 stupňoch voľnosti.

Test nulovosti prvkov korelačnej matice

Testujeme nulovú hypotézu, že $\rho_{ij}(k) = 0$ pre $|k| > q$ pre zvolené q (tj. skúmaná rada je postupnosť MA rádu q) proti alternatívnej hypotéze $\rho_{ij}(k) \neq 0$ pre

nejaké $|k| > q$. Potom vzhľadom k asymptoticky normálnemu rozdeleniu $\hat{\rho}_{ij}(k)$ zamietame na hladine α nulovú hypotézu v prospech alternatívnej ak

$$\frac{|\hat{\rho}_{ij}(k)|}{\hat{\sigma}_{ij}(k)} > u_{1-\frac{\alpha}{2}} \text{ pre nejaké } k > 0,$$

kde u_β značí β -kvantil štandardného normálneho rozdelenia a

$$\hat{\sigma}_{ij}(k) = \frac{1}{\sqrt{n-k}} \left(1 + 2 \sum_{s=1}^k \hat{\rho}_{ii}(s) \hat{\rho}_{jj}(s)\right)^{\frac{1}{2}}.$$

POZNÁMKA. Pre $\alpha = 0.05$ je $u_{1-\frac{\alpha}{2}} \doteq 2$, preto sa často pri testovaní nulovosti prvkov korelačnej matice porovnáva $|\hat{\rho}_{ij}(k)|$ s dvojnásobkom smerodatnej odchýlky $\hat{\sigma}_{ij}(k)$.

Test nulovosti korelačnej matice

S využitím chí-kvadrát štatistiky $R(k)$ sa dá testovať nulovosť celej korelačnej matice. Nulovosť zamietame ak

$$R(k) > \chi_{m^2}^2(1 - \alpha),$$

kde $\chi_{m^2}^2(\beta)$ značí β -kvantil chí-kvadrát rozdelenia o m^2 stupňoch voľnosti. Do $R(k)$ dosádzame v tomto prípade odhady $\hat{\sigma}_{ij}(k)$.

K identifikácii modelu AR(p) je potrebné zaviesť mnohorozmernú analógiu parciálnej autokorelačnej funkcie. To sa dá spraviť dvojakým spôsobom.

2.3.2. Parciálna autoregresná maticová funkcia.

DEFINÍCIA. Nech $\{\mathbf{X}_t\}$ je slabo stacionárna náhodná postupnosť, uvažujme mnohorozmerný lineárny regresný model

$$\mathbf{X}_{t+k} = \Phi_{k,1} \mathbf{X}_{t+k-1} + \Phi_{k,2} \mathbf{X}_{t+k-2} + \dots + \Phi_{k,k} \mathbf{X}_t + \mathbf{e}_{k,t+k}$$

kde \mathbf{X}_{t+k} je vysvetľovaná, $\mathbf{X}_{t+k-1}, \mathbf{X}_{t+k-2}, \dots, \mathbf{X}_t$ sú vysvetľujúce premenné a $\mathbf{e}_{k,t+k}$ je reziduálny vektor. Potom sa funkcia $\zeta(k) = \Phi_{k,k}$ definovaná pre $k = 1, 2, \dots$ nazýva *parciálna autoregresná maticová funkcia* náhodnej postupnosti $\{\mathbf{X}_t\}$.

POZNÁMKA. Zavedme nasledujúce označenie matíc:

$$\mathbf{A}(k) = \begin{pmatrix} \mathbf{\Gamma}(0) & \mathbf{\Gamma}(1)^T & \dots & \mathbf{\Gamma}(k-2)^T \\ \mathbf{\Gamma}(1) & \mathbf{\Gamma}(0) & \dots & \mathbf{\Gamma}(k-3)^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Gamma}(k-2) & \mathbf{\Gamma}(k-3) & \dots & \mathbf{\Gamma}(0) \end{pmatrix},$$

$$b(k) = \begin{pmatrix} \mathbf{\Gamma}(k-1)^T \\ \mathbf{\Gamma}(k-2)^T \\ \vdots \\ \mathbf{\Gamma}(1)^T \end{pmatrix}, c(k) = \begin{pmatrix} \mathbf{\Gamma}(1) \\ \mathbf{\Gamma}(2) \\ \vdots \\ \mathbf{\Gamma}(k-1) \end{pmatrix}.$$

Potom je

$$\zeta(k) = \begin{cases} \mathbf{\Gamma}(1)^T \mathbf{\Gamma}(0)^{-1}, & k = 1, \\ (\mathbf{\Gamma}(k)^T - c(k)^T (\mathbf{A}(k))^{-1} b(k)) (\mathbf{\Gamma}(0)^T - b(k)^T (\mathbf{A}(k))^{-1} b(k))^{-1}, & k > 1. \end{cases}$$

Dá sa ukázať, že parciálna autoregresná maticová funkcia $\zeta(k)$ postupnosti AR(p) je nulová pre $k > p$. Navyiac platí $\zeta(p) = \Phi_p$.

DEFINÍCIA. Nech časová rada $\{\mathbf{x}_t, t = 1, 2, \dots, n\}, n \in \mathbb{N}$ je realizáciou náhodnej postupnosti $\{\mathbf{X}_t, t \in \mathbb{N}\}$. Potom sa matica $\hat{\zeta}(k)$, vypočítaná podľa poznámky v 2.3.2, kde sa namiesto $\Gamma(s)$ pre $s = 0, 1, \dots, k-1$ dosadí matica $\hat{\Gamma}(s)$ s prvkami

$$\hat{\gamma}_{ij}(s) = \frac{1}{n} \sum_{t=1}^{n-s} (x_{t,i} - \bar{x}_i)(x_{t+s,j} - \bar{x}_j)$$

nazýva *výberová parciálna autoregresná matica*.

2.3.3. Parciálna korelačná maticová funkcia.

DEFINÍCIA. Nech $\{\mathbf{X}_t\}$ je slabo stacionárna náhodná postupnosť, $t \in \mathbb{Z}, k \in \mathbb{N}$

$$\mathbf{u}_{k-1,t+k} = \mathbf{X}_{t+k} - \alpha_{k-1,1}\mathbf{X}_{t+k-1} - \alpha_{k-1,2}\mathbf{X}_{t+k-2} - \dots - \alpha_{k-1,k-1}\mathbf{X}_{t+1},$$

$$\mathbf{v}_{k-1,t} = \mathbf{X}_t - \beta_{k-1,1}\mathbf{X}_{t+1} - \beta_{k-1,2}\mathbf{X}_{t+2} - \dots - \beta_{k-1,k-1}\mathbf{X}_{t+k-1}$$

kde $\alpha_{k-1,s}\beta_{k-1,s}$ pro $s = 1, 2, \dots, k-1$ sú matice minimalizujúce $E(\mathbf{u}_{k-1,t+k})^2$, resp. $E(\mathbf{v}_{k-1,t})^2$ a pre ktoré značíme

$$\mathbf{V}_u(k) = \text{var}(\mathbf{u}_{k-1,t+k}),$$

$$\mathbf{V}_v(k) = \text{var}(\mathbf{v}_{k-1,t}),$$

$$\mathbf{V}_{vu}(k) = \text{cov}(\mathbf{u}_{k-1,t+k}\mathbf{v}_{k-1,t}).$$

Nech $\mathbf{D}_v(k)$ (resp. $\mathbf{D}_u(k)$) je diagonálna matica s odmocninami diagonálnych prvkov matice $\mathbf{V}_v(k)$ na diagonále (resp. $\mathbf{V}_u(k)$). Potom sa funkcia

$\xi(k) = (\mathbf{D}_v(k))^{-1}\mathbf{V}_{vu}(k)\mathbf{D}_u(k)^{-1}$ definovaná pre k kladné nazýva *parciálna korelačná maticová funkcia* náhodnej postupnosti $\{\mathbf{X}_t\}$.

POZNÁMKA. $\mathbf{V}_u(k), \mathbf{V}_v(k), \mathbf{V}_{vu}(k)$ môžeme vyjadriť pomocou kovariančnej maticovej funkcie $\Gamma(s)$, $s = 0, 1, \dots, k$. Príslušné vzorce sú uvedené v knihe [2].

Dá sa ukázať, že ak je $\{\mathbf{X}_t\}$ slabo stacionárna náhodná postupnosť, $\zeta(k)$ jej parciálna autoregresná maticová funkcia a $\xi(k)$ jej parciálna korelačná maticová funkcia. Potom je $\zeta(k) = 0$ práve pre také k , pre ktoré je $\xi(k) = 0$. To znamená, že parciálna korelačná maticová funkcia $\xi(k)$ postupnosti AR(p) je nulová pre $k > p$.

DEFINÍCIA. Nech časová rada $\{\mathbf{x}_t, t = 1, 2, \dots, n\}$ je realizáciou náhodnej postupnosti $\{\mathbf{X}_t, t \in \mathbb{N}\}$. Potom sa matica $\hat{\xi}(k)$ vypočítaná podľa [2] na strane 359-361, pri ktorej výpočte sa namiesto $\Gamma(s)$ pre $s = 0, 1, \dots, k$ použije matica $\hat{\Gamma}(s)$ s prvkami

$$\hat{\gamma}_{ij}(s) = \frac{1}{n} \sum_{t=1}^{n-s} (x_{t,i} - \bar{x}_i)(x_{t+s,j} - \bar{x}_j),$$

nazýva *výberová parciálna korelačná matica*.

Dá sa ukázať, že pokiaľ je časová rada $\{\mathbf{x}_t, t = 1, 2, \dots, n\}, n \in N$ realizáciou postupnosti AR(p), potom pre $k > p$ a pre veľké n má $\hat{\xi}_{ij}(k)$ asymptoticky normálne rozdelenie s nulovou strednou hodnotou a rozptylom $\frac{1}{n-k}$.

Naviac pre $k > p$ má štatistika

$$X(k) = (n-k) \sum_{i=1}^m \sum_{j=1}^m (\hat{\xi}_{ij}(k))^2$$

asymptoticky chí-kvadrát rozdelenie o m^2 stupňoch voľnosti.

Test nulovosti prvkov parciálnej korelačnej matice

Pri testovaní nulovosti prvkov parciálnej korelačnej matice sa využíva ich asymptoticky normálne rozdelenie. Nulovosť zamietame na hladine α , pokiaľ

$$|\hat{\xi}_{ij}(k)| > u_{1-\frac{\alpha}{2}} \frac{1}{\sqrt{n-k}}, \text{ resp. } |\hat{\xi}_{ij}(k)| > \frac{2}{\sqrt{n-k}} \text{ pre } \alpha = 0.05.$$

Test nulovosti parciálnej korelačnej matice

Test nulovosti celej parciálnej korelačnej matice je založený na asymptotickom rozdelení štatistiky $X(k)$. Nulovosť zamietame na hladine α , pokiaľ

$$X(k) > \chi_{m^2}^2(1 - \alpha).$$

Zhrnutie voľby rádu modelu

Napriek tomu, že výpočet výberovej korelačnej maticovej funkcie a výberovej parciálnej autoregresnej alebo parciálnej korelačnej maticovej funkcie je potrebný k odhadu rádu postupnosti MA(q) resp. AR(p), nedávajú návod ako voliť rády zmiešaného modelu ARMA(p,q).

U mnohorozmerných časových rád s menej zložkami je možné skúmať korelačnú štruktúru po zložkách a riadiť sa pravidlami používanými pri identifikácii modelov jednorozmerných rád.

Druhou možnosťou je použitie štatistických rozhodovacích kritérií, ktoré umožňujú automatizovať postup identifikácie modelu.

Rozhodovacie kritériá

Akaikeho informačné kritérium (AIC)

$$AIC(p, q) = \log(\det(\hat{\Sigma}_{p,q})) + \frac{2m^2(p+q)}{n},$$

Baysové informačné kritérium (BIC) tvaru

$$BIC(p, q) = \log(\det(\hat{\Sigma}_{p,q})) + 2m^2(p+q) \frac{\log n}{n},$$

kde p a q sú rády modelu a $\hat{\Sigma}_{p,q}$ príslušný odhad rozptylovej matice bieleho šumu. Za optimálny model je považovaný model s minimálnym AIC (resp. BIC).

2.4. Verifikácia modelu**Test nulovosti rezíduí**

Test nulovosti rezíduí je založený na skutočnosti, že za nulovej hypotézy sú rezídua \hat{y}_t realizáciami bieleho šumu, u ktorého predpokladáme nulovú strednú hodnotu, nekorelovanosť pre rôzne t a pre ktorý máme odhad $\hat{\Sigma}$ rozptylovej matice Σ (ich prvky označme $\hat{\sigma}_{ij}$, resp. σ_{ij}). Za predpokladu normality a použitím centrálnej limitnej vety to znamená, že výberové priemery $\bar{y}_i = \frac{1}{n-r} \sum_{t=r+1}^n \hat{y}_{i,t}$, $i = 1, 2, \dots, m$ majú za nulovej hypotézy asymptoticky normálne rozdelenie s nulovou strednou hodnotou a rozptylom $\frac{\sigma_{ii}}{n-r}$.

Nulovú hypotézu zamietame na hladine α v prípade, že pre nejaké i je

$$|\bar{y}_i| > u_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}_{ii}}{n-r}}, \text{ resp. } |\bar{y}_i| > 2\sqrt{\frac{\hat{\sigma}_{ii}}{n-r}} \text{ pre } \alpha = 0.05.$$

Pokiaľ navyše predpokladáme, že biely šum má normálne rozdelenie (takýto predpoklad napríklad umožní pre odhad parametrov použiť metódu maximálnej vierohodnosti), je možné testovať priamo nulovosť jednotlivých rezíduí, konkrétne nulovosť $y_{i,t}$ zamietame na hladine α , pokiaľ

$$|\hat{y}_{i,t}| > u_{1-\frac{\alpha}{2}} \sqrt{\sigma_{ii}}, \text{ resp. } |\hat{y}_{i,t}| > 2\sqrt{\sigma_{ii}} \text{ pre } \alpha = 0.05.$$

Test nekorelovanosti rezíduí

K testu nekorelovanosti rezíduí sa použije výberová korelačná matica vypočítaná pre radu $\{\hat{\mathbf{y}}_t\}$, pretože za nulovej hypotézy sú jednotlivé rezídua realizáciou bieleho šumu, ktorého korelačná maticová funkcia $\rho(k)$ je nulová pre $k \neq 0$. Dá sa ukázať, že $\hat{\rho}_{ij}(k)$ majú asymptoticky normálne rozdelenie s nulovou strednou hodnotou a rozptylom $\frac{1}{n-k}$.

Nulovosť nejakého $\rho_{ij}(k)$ zamietame na hladine α , pokiaľ

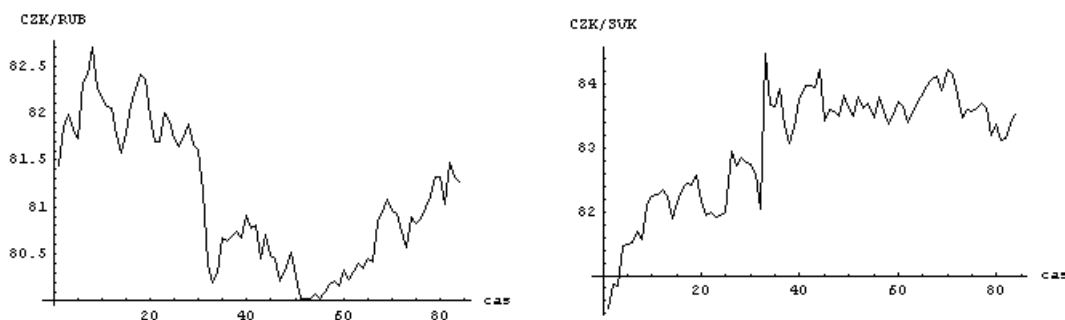
$$|\hat{\rho}_{ij}(k)| > u_{1-\frac{\alpha}{2}} \frac{1}{\sqrt{n-k}}, \text{ resp. } |\hat{\rho}_{ij}(k)| > \frac{2}{\sqrt{n-k}} \text{ pre } \alpha = 0.05.$$

Je možné tolerovať niekoľko ojedinelých výnimiek.

2.5. Analýza dát

V praktickej časti venovanej mnohorozmerným radám budem analyzovať denný vývoj ruského rubľa a slovenskej koruny oproti českej korune. Dáta sú získané z webových stránok Českej národnej banky a jedná sa o rady denných kurzov od 1.2.2007 do 1.6.2007, celkovo 84 pozorovaní.

V obrázku 1 vidíme priebeh oboch časových rád.



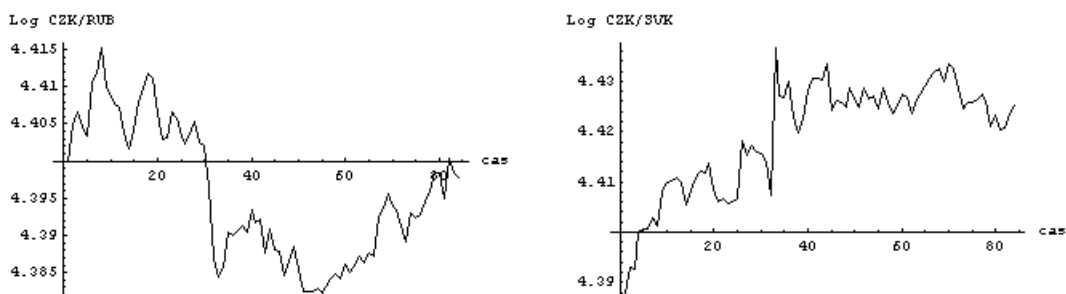
OBR. 1. Rady menových kurzov

Transformácia časovej rady

Pred analýzou časovej rady je nutné túto radu transformovať na radu stacionárnu a centrovanú.

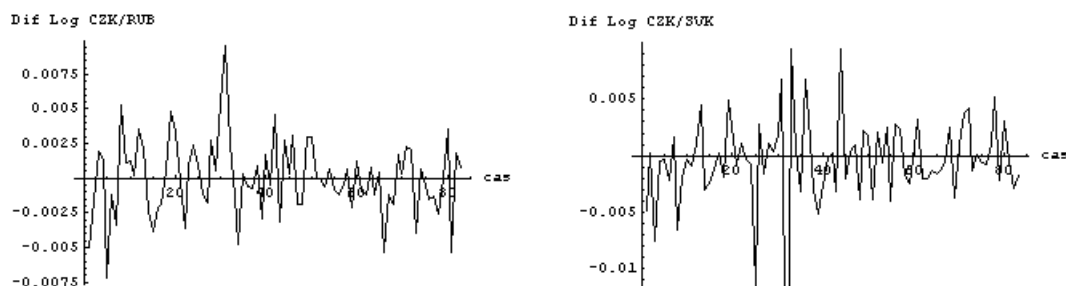
Prvou fázou je stacionarizácia rozptylu rady. Existuje viac kritérií, ktoré umožňujú jej stacionarizáciu. Najčastejšie sa však používa logaritmovanie.

Priebeh logaritmovaných rád ukazuje obrázok 2.



OBR. 2. Logaritmované rady

Stacionarizáciu v strednej hodnote, to je centrovanie rady, zabezpečíme prejdением k rade diferencií. Uprednostňujú sa diferenciie nižších rádov. Na logaritmované rady teda boli ďalej aplikované 1.diferencie, výsledok vidíme v obrázku 3.

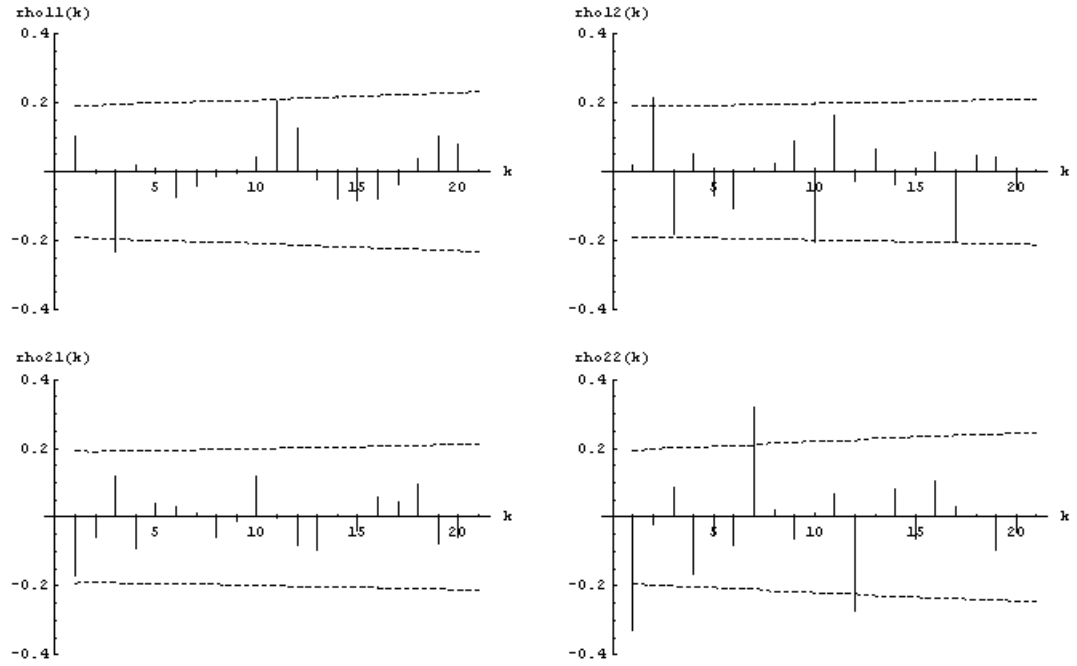


OBR. 3. Rady diferencií logaritmov

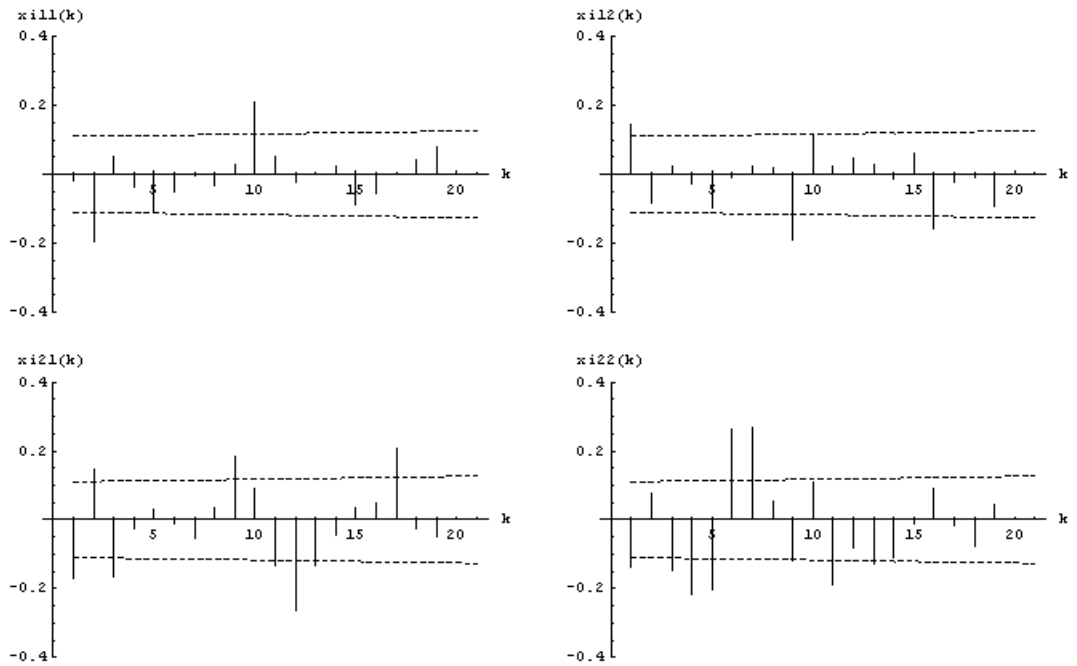
Stacionarita v strednej hodnote je uspokojivá, dokonalej stacionarity v rozptyle sa nepodarilo dosiahnuť.

Identifikácia ARMA modelu

Na identifikáciu použijeme hodnoty výberovej korelačnej matice $\hat{\rho}(k)$ a výberovej parciálnej korelačnej matice $\hat{\xi}(k)$ pre $k = 1, 2, \dots, 20$. Na grafoch sú znázornené spolu s pásom tvoreným ich odhadnutými smerodatnými odchýlkami. Pre test nulovosti jednotlivých prvkov korelačnej matice na hladine 5 percent sa používa porovnanie s dvojnásobkom odhadnutej smerodatnej odchýlky, ale vzhľadom k tomu, že účelom tohto skúmania korelačnej maticovej funkcie není ani tak samotné testovanie nulovosti ako skôr odhadnutie možných rádov modelu prostredníctvom odhalenia všetkých potenciálne nenulových prvkov, postup sprísňime práve prechodom k samotnej odhadnutej smerodatnej odchýlke. Tým je zaručené, že budeme ľahšie upozornení na nenulovosť niektorého prvku korelačnej alebo parciálnej korelačnej matice.

OBR. 4. Odhadnuté hodnoty prvkov $\rho_{ij}(k)$ výberovej korelačnej matice

Z grafov korelačnej maticovej funkcie vidíme, že mimo vyznačený pás sa dostali len 3 hodnoty v grafe $\hat{\rho}_{22}(k)$ a aj tieto by nepresiahli dvojnásobok smerodatnej odchýlky. Nejde teda identifikovať MA model.

OBR. 5. Odhadnuté hodnoty prvkov $\xi_{ij}(k)$ výberovej parciálnej korelačnej matice

V priebehoch grafov $\hat{\xi}_{21}(k)$ a $\hat{\xi}_{22}(k)$ dochádza k výraznejšiemu prekročeniu vyznačených hraníc u niektorých hodnôt.

Skúmané boli ARMA modely do rádu 5. Pre zúženie výberu modelov použijeme Akaikeho informačné kritérium. Za najvhodnejší model bol zvolený model ARMA(1,2).

Hodnota AIC je v tomto prípade -22.9023 . Na základe využitia metódy maximálnej vierohodnosti boli zistené odhady parametrov modelu. Ak odhadnuté parametre nepresahovali odhad svojej smerodatnej odchýlky, tak boli nahradené nulami.

	ARMA(1,2)	
$\hat{\Phi}_1$	0	0.9654353
	0.3112664	0
$\hat{\Theta}_1$	0	-1.0283223
	-0.6252920	-0.6466818
$\hat{\Theta}_2$	0.3999107	0.3709635
	0	0
$\hat{\Sigma}$	0.0000061	0.0000007
	0.0000007	0.0000138

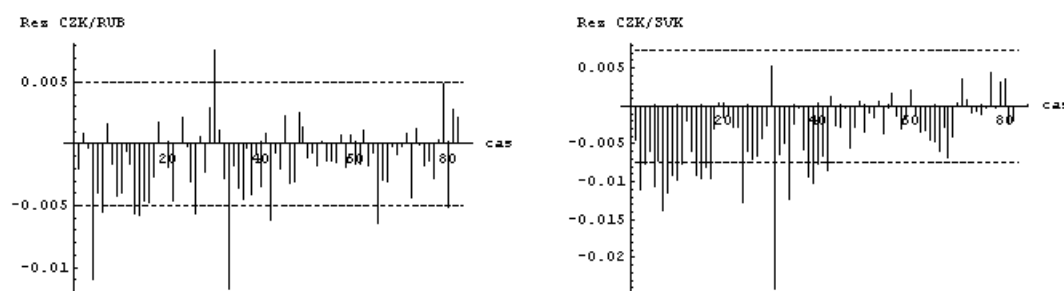
Verifikácia ARMA modelu

Budeme overovať správnosť voľby rádu modelu, správnosť odhadu parametrov a oprávnenosť predpokladu normality bieleho šumu, ktorý je obsiahnutý v použití metódy maximálnej vierohodnosti pre odhad parametrov. K posúdeniu nulovosti strednej hodnoty rezíduí použijeme najprv test nulovosti porovnávajúci absolútnu hodnotu priemeru s dvojnásobkom jeho smerodatnej odchýlky. V tabuľke sú obsiahnuté priemery zložiek rezíduí a dolné medze testov ich nulovosti na hladine 5 percent.

	ARMA(1,2)
\bar{y}_1	-0.0015945
$2\hat{\sigma}_{\bar{y}_1}$	0.0005444
\bar{y}_2	-0.0041091
$2\hat{\sigma}_{\bar{y}_2}$	0.0008177

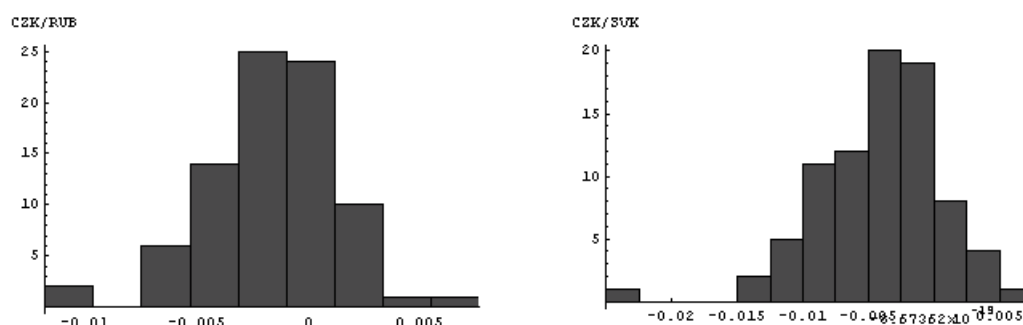
Z tabuľky vidíme, že $\bar{y}_j > 2\hat{\sigma}_{\bar{y}_j}$, $j = 1, 2$ a teda zamietame nulovosť strednej hodnoty oboch zložiek vektoru rezíduí.

V obrázku 6 vidíme, ktoré konkrétne rezídua v absolútnej hodnote prekročili dvojnásobok svojej smerodatnej odchýlky.



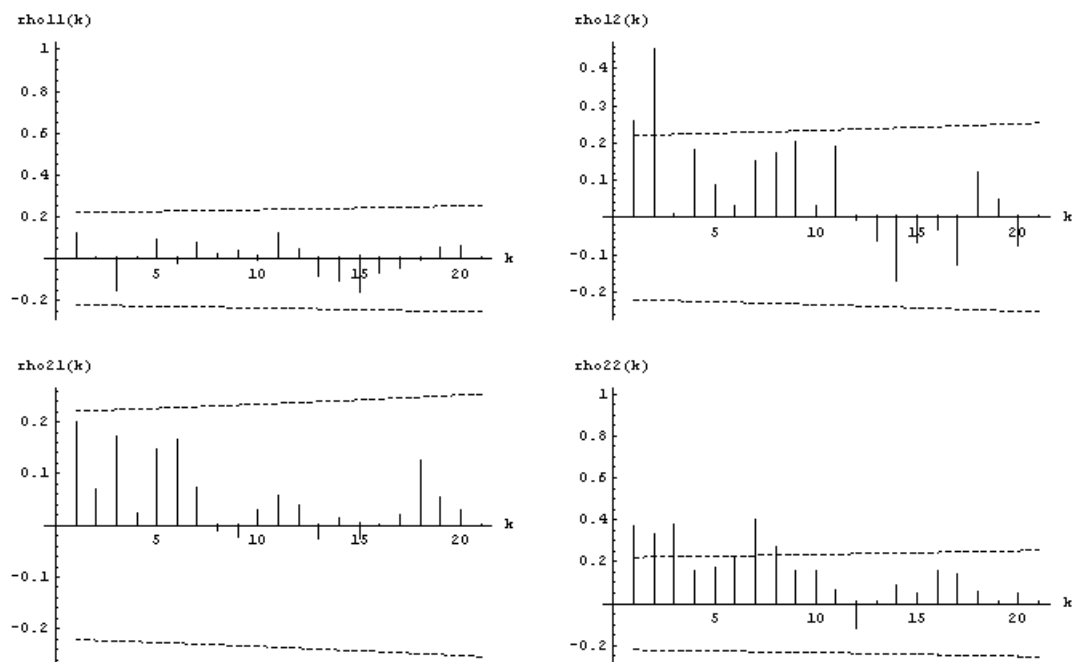
OBR. 6. Rezídua modelu ARMA(1,2)

Normalita rezíduí bola posudzovaná na základe porovnania histogramov rezíduí s tvarom hustoty normálneho rozdelenia. Vidíme, že histogramy v okolí nuly vykazujú približne priebeh normálneho rozdelenia, ak odhliadneme od niekoľkých extrémne malých odľahlých hodnôt.



OBR. 7. Histogramy rezíduí

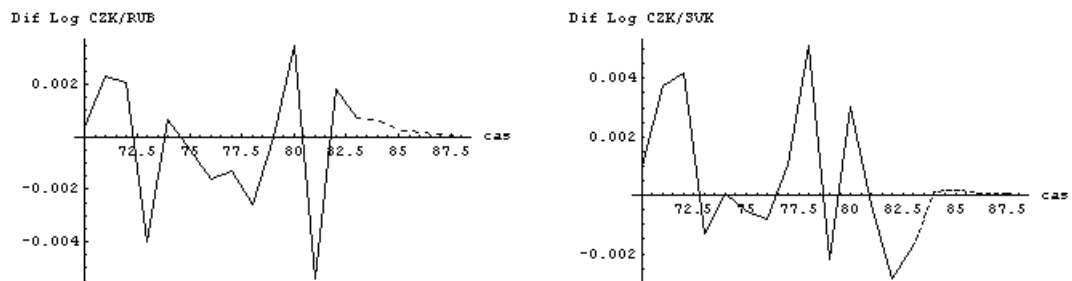
Posúdenie korelovanosti rezíduí je analogické ako v prípade analýzy korelačnej maticovej funkcie skúmanej rady. Pásky v grafe sú určené dvojnásobkom odhadnutých smerodatných odchýliek jednotlivých korelácií. Za výrazne odlišnú od nuly je možné považovať len hodnotu $\hat{\rho}_{12}(2)$. Napriek ojedinelým hodnotám v $\hat{\rho}_{22}(k)$, ktoré nevýrazne prekročili dvojnásobok svojej smerodatnej odchýlky, môžeme konštatovať, že rezídua sú takmer nekorelované.



OBR. 8. Odhadnuté hodnoty prvkov $\rho_{ij}(k)$ výberovej korelačnej matice rezíduí

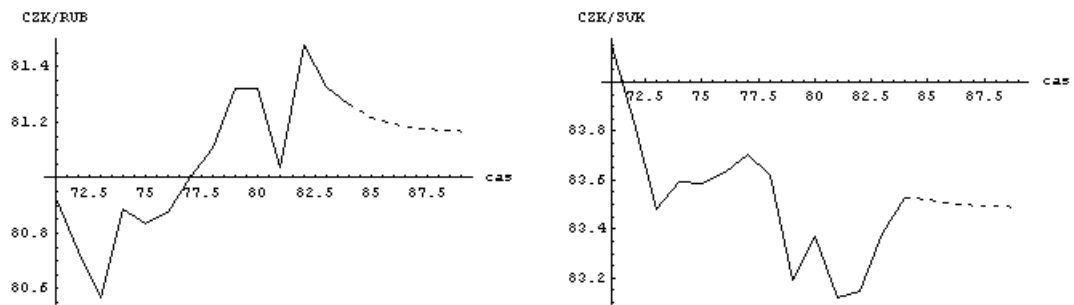
Konštrukcia predpovedí a spätná transformácia

Zvolený model ARMA(1,2) použijeme ku konštrukcii 5 predpovedí. V grafe je znázornených posledných 13 pozorovaných hodnôt a prvých 5 predpovedí.



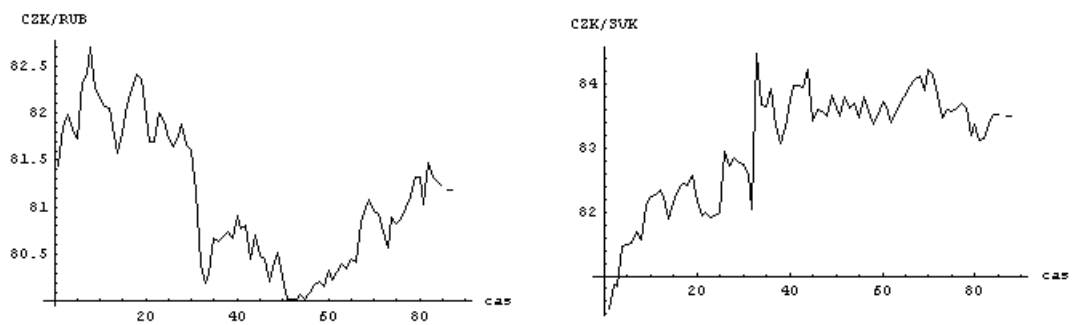
OBR. 9. Transformovaná rada s predpoveďami

Priebeh skráteného grafu po spätnej transformácii ukazuje obrázok 10.



OBR. 10. Pôvodná rada s predpoveďami

Pre úplnosť doplnený kompletný graf aj s predpoveďami.



OBR. 11. Kompletná pôvodná rada s predpoveďami

Záver

Napriek faktu, že vo fáze verifikácie sa ukázala ako sporná nulová stredná hodnota a vďaka niekoľkým extrémnym hodnotám aj normalita rezíduí, použili sme model ARMA(1,2) navrhnutý na základe minimalizácie AIC pre konštrukciu predpovedí. Vo vývoji kurzu CZK/RUB je predpovedané pokračovanie v tendencii spomaľujúceho sa poklesu, predpoveď vývoja kurzu CZK/SVK dáva menej uspokojivú informáciu.

Záver

Cieľom tejto práce bolo popísať teóriu jednorozmerných a mnohorozmerných modelov časových rád a aplikovať ju s použitím dostupného softwaru na konkrétne dáta z finančných trhov.

V teoretickej časti je popísané exponenciálne vyrovňovanie a Boxova-Jenkinsova metodológia, to jest jednorozmerné a mnohorozmerné ARMA modely. Sú tu obsiahnuté všetky definície potrebné na analýzu konkrétnych časových rád, od problému identifikácie modelu až po predpovede budúcich hodnôt.

V praktickej časti sú spracovávané dve časové rady. Ide o radu menových kurzov švajčiarskeho franku ku českej korune spracovávanú programom NCSS a dvojrozmernú radu menových kurzov ruského rubľa a slovenskej koruny ku českej korune spracovávanú programom vytvoreným v práci [4] v systéme Mathematica. V oboch prípadoch boli dáta stiahnuté z webových stránok Českej národnej banky. Išlo o rady denných kurzov z prvej polovice roka 2007. Použité programy (rada.nb, franky.S0) sú priložené na CD.

V súčasnosti je vďaka rozvoju výpočtovej techniky možné na analýzu časových rád použiť aj zložitejšie nelineárne modely. Napriek tomu si lineárne ARMA modely tiež zaslúžia našu pozornosť. Predstavujú totiž výhodný kompromis v zmysle zachovania nízkej výpočtovej náročnosti a pomerne dobrými schopnosťami popisu skúmaných dát. Táto ich vlastnosť je cenená hlavne v prípade mnohorozmerných rád, kde hlavne pri vyšších rozmeroch dochádza k neúmernému rastu výpočtovej náročnosti.

Literatúra

- [1] Cipra T.: *Analýza časových řad s aplikacemi v ekonomii*, SNTL, Praha, 1986.
- [2] Wei W.W.S.: *Time Series Analysis. Univariate and Multivariate Methods*, Addison-Wesley, New York, 1990.
- [3] Mathematica: *Time Series Pack. Reference and User's Guide*, Wolfram Research, 1995.
- [4] Hrba T.: *Diplomová práce*, Univerzita Karlova v Praze Matematicko-fyzikální fakulta, 2006.
- [5] www.cnb.cz